

From: Fredman, Jeffrey
Sent: Wednesday, November 14, 2001 12:57 PM
To: STIC-Biotech/ChemLib
Cc: Bugaisky, Gabriele
Subject: FW: 09/215435

PLEASE RUSH!

I Approve.

Jeff Fredman

-----Original Message-----

Fr m: Bugaisky, Gabriele
Sent: Wednesday, November 14, 2001 12:08 PM
To: Fredman, Jeffrey
Subject: 09/215435

Edward Hart
Technical Info Specialist
STIC / Biotech
CM 12C14 Tel: 305-9203

This is an old est case that needs further search:

please search a) polynt encoding all of SEQ ID NO:177 and the following portions of SEQ ID NO:177: aa 1-24 and 25-99.

b)polynt encoding all of SEQ ID NO:179 and the following portions of SEQ ID NO:179: aa 1-23 and 24-121

c)polynt encodingall of SEQ ID NO:225 and the following portions of SEQ ID NO:225:aa 1-22 and 23-227.

originally only specific nt sequences were searched, not any nt encoding the protein.

please rush, if possible, as I have to reopen AF.

thanks, gabi b.

Gabriele E. Bugaisky

- au 1653
- cm1-10d09; mailbox 9b01
- 308-4201

RECEIVED
NOV 14 2001
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 11/14/01
Date Completed: 11/15/01
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 9 → reverse to NA
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: EV
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-177 to: GenEmbl.* out_format : pfs

Date: Nov 15, 2001 12:38 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgml_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -QEMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435 -CGCML_1_11006 -NCFU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177

Query length: 99

Database: GenEmbl.*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 7698.940000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb.pr.9: HSDJ60101	+	341.00	638.54	2.5e-27	33458	! AL109656 Human DNA sequence f
gb.htg9: AC023971	+	341.00	625.34	1.4e-26	153788	! AC023971 Homo sapiens chromo
gb.in2: AF078161	+	141.00	249.48	1.2e-05	11008	! AF078161 Manduca sexta lacuni
gb.pat2: E12856	+	135.50	264.37	1.7e-06	555	! E12856 Synthetic DNA encoding
gb.om: BTU35442	+	134.00	254.72	6.0e-06	1198	! U35442 Bos taurus alpha1-micro
gb.in3: CBL37C3	+	133.00	221.83	0.0004	42521	! U64857 Caenorhabditis elegans
gb.htg1: AC006915	+	133.00	211.89	0.0014	134095	! AC006915 Caenorhabditis eleg
gb.htg6: AC001993	+	132.50	220.00	0.0005	80689	! AC001993 Drosophila melanoga
gb.in1: AC003053	+	132.50	215.29	0.0009	80689	! AC003053 Drosophila melanoga
gb.in1: AC005149	+	132.50	213.52	0.0012	98942	! AC005149 Drosophila melanoga
gb.in1: AE003579	+	132.00	203.80	0.0041	304383	! AE003579 Drosophila melanoga
gb.pat2: T63559	+	131.00	247.06	1.6e-05	1455	! I63559 Sequence 6 from patent
gb.om: MEELACR	+	130.00	255.84	5.2e-06	4191	! AJ000490 Macropus eugenii mR
gb.om: HSA1MICR	+	129.50	245.58	1.9e-05	1221	! X04494 Human mRNA for alpha-1
gb.pat2: T05217	+	129.50	245.51	1.9e-05	1232	! I05217 Sequence 7 from Patent
gb.pr.9: HSHCR	+	129.50	245.46	2.0e-05	1239	! X04225 Human mRNA for protein
gb.pr.9: HSA1MBG3	+	129.50	234.05	8.4e-05	4627	! X54818 Human gene for alpha-1
gb.pr.10: HUMITILC08	+	129.50	232.68	0.0001	5423	! M88249 Human inter-alpha-tryps
gb.htg19: AL137850	+	129.50	201.72	0.0053	193939	! AL137850 Homo sapiens chrom
gb.htg3: AC011161	+	129.50	201.43	0.0055	200532	! AC011161 Homo sapiens clone
gb.sv: SYNBOVTRV1	+	129.00	256.71	4.6e-06	301	! M31335 Bovine synthetic colostr
gb.in2: CBRG45N02	+	129.00	213.82	0.0011	42726	! AC084549 Caenorhabditis brig
gb.pat1: A31024	+	127.00	256.66	4.6e-06	191	! A31024 DNA fragment for bikunin
gb.pat1: A31025	+	127.00	256.66	4.6e-06	191	! A31025 DNA fragment for bikunin
gb.pat1: A19212	+	127.00	249.88	1.1e-05	418	! AL19212 Human kunitz type protea
gb.pat2: T29002	+	127.00	249.88	1.1e-05	418	! T29002 Sequence 17 from patent
gb.om: TVU34208	+	126.00	248.76	1.3e-05	378	! U34208 Trichosurus vulpecula ea
gb.r01: D50586	+	125.50	235.70	6.8e-05	1522	! D50586 Mus musculus mRNA for P
gb.pat2: E13092	+	124.00	243.55	2.5e-05	435	! E13092 DNA encoding a protease
gb.pr.5: AK024817	+	124.00	239.10	4.4e-05	728	! AK024817 Homo sapiens cDNA: FL
gb.pat2: T14875	+	124.00	236.53	6.1e-05	979	! T14875 Sequence 1 from patent
gb.pat2: T92685	+	124.00	236.53	6.1e-05	979	! T92685 Sequence 1 from patent
gb.pr.10: HUMHTFP	+	124.00	236.53	6.1e-05	979	! L27624 Homo sapiens tissue fact
gb.s02: G31809	+	124.00	235.20	7.3e-05	1142	! G31809 SW554035 Eric D. Green
gb.pr.10: HUMMP5	+	124.00	235.20	7.3e-05	1142	! D29992 Homo sapiens mRNA for P
gb.r01: MNAMP	+	124.00	234.92	7.6e-05	1180	! X68680 M.musculus AMP mRNA f
gb.pr.7: BC005330	+	124.00	234.73	7.7e-05	1203	! BC005330 Homo sapiens, tissue
gb.r01: MUSA1ATI	+	124.00	234.53	7.9e-05	1234	! D28812 Mouse mRNA for alpha-1
gb.pr.9: HSDJ44793	+	124.00	196.72	0.0101	97385	! AL050348 Human DNA sequence f
gb.om: OAO000163	+	123.50	250.49	1.0e-05	174	! U00163 Ovis aries secretory pro
gb.om: AF241828	+	123.50	234.54	7.9e-05	1098	! AF241828 Ovis aries kunitz dom

gb.pat1: AL19210 + 123.00 241.91 3.1e-05 418 ! AL19210 Human kunitz type pro
gb.pat2: I29001 + 123.00 241.91 3.1e-05 418 ! I29001 Sequence 15 from pate
gb.in3: CEW01F3 + 123.00 204.14 0.0039 32797 ! Z92815 Caenorhabditis eleg
gb.htg24: CEV44A6 + 123.00 184.26 0.0501 326074 ! Z98863 Caenorhabditis ele
seq_name: gb.pr.9: HSDJ60101

seq_documentation_block:

LOCUS HSDJ60101 33458 bp DNA PRI 04-APR-2001
DEFINITION Human DNA sequence from clone RP4-60101 on chromosome 20. Contains
the gene for a novel protein with Kunitz/Bovine pancreatic trypsin
inhibitor domain, the 5' end of a novel gene, part of the gene for
a novel protein with WAP-type (Whey Acidic Protein) 'four-disulfide
core' domain, ESTs and GSSs, complete sequence.

ACCESSION AL109656

VERSION AL109656.10 GI:10944958

KEYWORDS HTG; Kunitz; WAP.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 33458)

AUTHORS Laird G.

TITLE Direct Submission

JOURNAL Submitted (03-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

requests: clonerequest@sanger.ac.uk

COMMENT On Oct 21, 2000 this sequence version replaced gi-9795206.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accessions

numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr20

IMPORTANT: This sequence is not the entire insert of clone

RP4-60101 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone RP3-447F3 is at 33359 in this sequence.

The true right end of clone RP4-688G8 is at 100 in this sequence.

RP4-60101 is from the library RPCI-4 constructed by the group of

Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest.

location/Qualifiers

1..33458

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="20"

/clone="RP4-60101"

/clone_lib="RPCI-4"

3..225

repeat_region /note="L1M4 repeat: matches 3977..4203 of consensus"

225..428

repeat_region /note="L1M4 repeat: matches 4321..4525 of consensus"

429..696

repeat_region /note="AluJb repeat: matches 2..302 of consensus"

697..778

repeat_region	/note="L1M4 repeat: matches 4525. .4606 of consensus" 779. .1089	repeat_region	/note="L2 repeat: matches 2666. .2732 of consensus" 13202. .13331
repeat_region	/note="AluX repeat: matches 1. .311 of consensus" 1090. .1146	repeat_region	/note="AluSg/x repeat: matches 183. .312 of consensus" 13332. .15510
repeat_region	/note="L1M4 repeat: matches 4606. .4662 of consensus" 1165. .1293	repeat_region	/note="L1M1 repeat: matches 3916. .6171 of consensus" 15534. .15832
repeat_region	/note="L1M4 repeat: matches 1. .134 of consensus" 1351. .1730	repeat_region	/note="AluSg repeat: matches 1. .299 of consensus" 15846. .16746
repeat_region	/note="L1M5 repeat: matches 5005. .5402 of consensus" 1731. .2032	repeat_region	/note="L1M4 repeat: matches 3472. .4394 of consensus" 16739. .18594
repeat_region	/note="AluSx repeat: matches 1. .301 of consensus" 2056. .2547	repeat_region	/note="L1P11 repeat: matches 11. .1921 of consensus" 18595. .18888
repeat_region	/note="L1M5 repeat: matches 5374. .5865 of consensus" 2543. .2729	repeat_region	/note="AluSg repeat: matches 1. .295 of consensus" 18889. .20972
repeat_region	/note="L1M4/D repeat: matches 5428. .5609 of consensus" 2730. .3022	repeat_region	/note="L1P11 repeat: matches 1. .310 of consensus" 20973. .21280
repeat_region	/note="AluSg repeat: matches 1. .293 of consensus" 3023. .3181	repeat_region	/note="AluSg repeat: matches 1. .310 of consensus" 21281. .23911
repeat_region	/note="L1M4/D repeat: matches 5609. .5769 of consensus" 3361. .3895	repeat_region	/note="L1P11 repeat: matches 3456. .6164 of consensus" 23926. .24129
repeat_region	/note="L1 repeat: matches 3599. .4167 of consensus" 3986. .4281	repeat_region	/note="L1M4 repeat: matches 4390. .4613 of consensus" 24130. .24438
repeat_region	/note="L2 repeat: matches 534. .834 of consensus" 4526. .4828	repeat_region	/note="AluX repeat: matches 1. .310 of consensus" 24439. .24513
misc_feature	/note="AluSx repeat: matches 1. .300 of consensus" 5372. .5726	repeat_region	/note="L1M4 repeat: matches 4613. .4683 of consensus" 24790. .25008
repeat_region	/note="match: GSS: Em:AQ076126" 5394. .5470	repeat_region	/note="MIR repeat: matches 1. .242 of consensus" 25397. .28876
gene	/note="MIR repeat: matches 48. .142 of consensus" 7401. .7609	repeat_region	join(25397. .25528,26926. .27103,28675. .28876) /gene="dj60101.1"
CDS	/gene="dj60101.3" <7401. .>7609	repeat_region	/note="match: CDNA: Em:AK020315 match: ESTs: Em:AW303354"
	/note="possibly a pseudogene match: proteins: Sw:Q14508 Sw:P00993 Sw:Q28631 Sw:P14730 Tr:O44131 Wp:C08G9"	repeat_region	/evidence=not_experimental /product="dj60101.1 (novel protein with Kunitz/Bovine pancreatic trypsin inhibitor domain)"
	/codon_start=2 /evidence=not_experimental /product="dj60101.3 (novel protein with WAP-type (Whey Acidic Protein) 'four-disulfide core' domain)"	repeat_region	/protein_id="CAC36025.1"
	/protein_id="CAC36024.1"	repeat_region	/db_xref="GI:13559238"
	/db_xref="GI:13559237"	repeat_region	/translation="AACCTMKSAKGLFLRRFFCSLNTLLGGVKNKIAKICGDLKD PCKLDMFGSCYEVHFRYFNRTSKRCETVFCSCGNLNFNFKIEREACVACVAKYK PR"
mRNA	/translation="PFGSSGLFSLSEHYSSLLVRRPFSYLCVPEVILPPPCISAPEN CTHLCTMQEDCEKGFQCCSSFCGIV" complement(7623. .7780) /genes="dj68868.4"	repeat_region	26066. .26228
	/note="Continues in Em:AL031671 as dj68868.4 match: ESTs: Em:AI808149 Em:AI808159"	repeat_region	/note="MIR repeat: matches 53. .205 of consensus" 27253. .27579
gene	/evidence=not_experimental /product="dj60101.2 (a putative novel protein)" complement(7623. .7780) /genes="dj68868.4"	repeat_region	/note="AluSp repeat: matches 1. .303 of consensus" 27649. .27728
repeat_region	/note="L2 repeat: matches 2596. .2670 of consensus" 8352. .8422	alignment_scores:	Quality: 341.00 Length: 60
repeat_region	8434. .8656	Ratio: 5.683 Gaps: 0	
repeat_region	/note="MIR repeat: matches 9. .260 of consensus" 9290. .9419	Percent Similarity: 100.000 Percent Identity: 100.000	
misc_feature	/note="MIR repeat: matches 47. .183 of consensus" complement(9433. .9899)	alignment_block:	
repeat_region	/note="match: GSS: Em:AQ476416" 9538. .9830	US-09-215-435-177 x HSDJ60101	
misc_feature	/note="AluSg repeat: matches 1. .293 of consensus" 9923. .10651	Align seg 1/1 to: HSDJ60101 from: 1 to: 33458	
repeat_region	/note="match: GSS: Em:AQ478366" 10453. .10647		
misc_feature	/note="MER3 repeat: matches 1. .204 of consensus" complement(10590. .10982)		
repeat_region	/note="match: GSS: Em:AQ223524" 11007. .11080		
repeat_region	/note="MER69 repeat: matches 2. .77 of consensus" 11529. .11954		
repeat_region	/note="MER69 repeat: matches 2070. .2511 of consensus" 13092. .13158		

```

39 AspProCysLysLeuAspMetAsnPhcGlySerCysTyrGluValHisph 55
|||||
26925 GATCCCTGCAAAATGACATGAATTTTGAAGCTGCTATGAAGTTCACTT 26974
|||||
55 eATgTYrPhETyrAsnArqThrSerLysArgCysGluThrPhcValPhes 72
|||||
26975 TAGATATTTCTACACAGAACCTCCAAAAGATGTGAACCTTTTGCTTCTT 27024
|||||
72 eTGlyCysAsnGlyAsnLeuAsnAsnPhcLysLeuLysIleGluArgGlu 88
|||||
27025 CCGGCTGTAATGCAACCTTAACAACCTTAAAGCTTAAATAAGAACGTGAA 27074
|||||
89 ValAlaCysValAlaLysTyrLysProPro 98
|||||
27075 GTAGCCTGTGTTGCCAAAATACAAACCCG 27104
|||||

```

seq_name: gb_htg9:AC023971

seq_documentation_block:
LOCUS AC023971 153788 bp DNA HTG 07-APR-2000
DEFINITION Homo sapiens chromosome 20 clone RP11-770N8 map 20, WORKING DRAFT
SEQUENCE, 13 unordered pieces.

ACCESSION AC023971
VERSION AC023971.2 GI:7523758
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 153788)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 20, clone RP11-770N8

REFERENCE 2 (bases 1 to 153788)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhalil,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepli,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArillano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginder,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrelli,A.,
Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

TITLE Direct Submission
JOURNAL Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Apr 7, 2000 this sequence version replaced gi:7008875.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6573
Center clone name: 770.N.8
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 145038 bases at least Q40
Consensus quality: 148490 bases at least Q30
Consensus quality: 150621 bases at least Q20
Insert size: 150000; agarose-fp
Insert size: 152588; sum-of-contigs
Quality coverage: 5.2 in Q20 bases; agarose-fp
Quality coverage: 5.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1333: contig of 1333 bp in length
* 1334 1433: gap of 100 bp
* 1434 3969: contig of 2536 bp in length
* 3970 4069: gap of 100 bp
* 4070 7211: contig of 3142 bp in length
* 7212 7311: gap of 100 bp
* 7312 13518: contig of 6207 bp in length
* 13519 13618: gap of 100 bp
* 13619 20876: contig of 7258 bp in length
* 20877 20976: gap of 100 bp
* 20977 28838: contig of 7862 bp in length
* 28839 28938: gap of 100 bp
* 28939 37590: contig of 8652 bp in length
* 37591 37690: gap of 100 bp
* 37691 45814: contig of 8124 bp in length
* 45815 45914: gap of 100 bp
* 45915 62601: contig of 16687 bp in length
* 62602 62701: gap of 100 bp
* 62702 81817: contig of 19116 bp in length
* 81818 81917: gap of 100 bp
* 81918 100273: contig of 18356 bp in length
* 100274 100373: gap of 100 bp
* 100374 125266: contig of 24893 bp in length
* 125267 125366: gap of 100 bp
* 125367 153788: contig of 28422 bp in length.

FEATURES

source
1..153788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/chromosome="20"
/clone="RP11-770N8"
/clone.lib="RPC1-11 Human Male BAC"
1..1333
/note="assembly_fragment"
1434..3969
/note="assembly_fragment"
4070..7211
/note="assembly_fragment"
7312..13518
/note="assembly_fragment"
13619..20876
/note="assembly_fragment"
20977..28838
/note="assembly_fragment"
28939..37590
/note="assembly_fragment"
37691..45814
/note="assembly_fragment"
clone_end:SP6
vector_side:right
45915..62601
/note="assembly_fragment"
62702..81817
/note="assembly_fragment"
clone_end:T7

misc_feature
1434..3969
/note="assembly_fragment"
4070..7211
/note="assembly_fragment"
7312..13518
/note="assembly_fragment"
13619..20876
/note="assembly_fragment"
20977..28838
/note="assembly_fragment"
28939..37590
/note="assembly_fragment"
37691..45814
/note="assembly_fragment"
clone_end:SP6
vector_side:right
45915..62601
/note="assembly_fragment"
62702..81817
/note="assembly_fragment"
clone_end:T7

Thu Nov 15 10:51:52 2001

```
/db_xref="GI:6164595"
/translation="MVLNLRSLVAVIVLSNLCIAWTASRHHYTHNVSGHRHRRQG
KGLYLSSSVYIPGGEGTGWGDSDTSCPTSCRTGGGVASQKRICLKPGDQPCQTGGD
TKYFSCQDQCPGDFDRAEQSDYDDVEYRNKIKWPKYRGNPCRLNMGHGER
FYRHLQVLDGTGRCNDSDVDVNGTQCPQGDMLGNARCKRKYRNGDKNCYIT
TNGVLDQDLIKYNDMLLIPGATITIVIMEVPSNNYLALRAKNKTYILNGDYLDF
PRSMIAGALWFYERSQGFAPDKLRCGLPTTEPLYLILLQSVANGVLEYEYSLPTE
RAPRPTQNYTWHEHTEFCSATCGGQFOTRNVTQTMOLLELVDDTLCDGELKPVTNQ
TCNNVPPQVWGEWGSCHRGSDGTOSRELOCEKLI VNGFRSIVSOKECFDLLGP
KPERKCNVNICPTWFTGPMKPCDLTLYGEGKQVQVYCKNGRRIDVLDSECTD
ERPESEQCMHSEERTDWASEWSDGNCSTMRIRIACTYIDKLVDDSPCAHHP
LPVLQEKDAATLPQVQWYATQWSKCEGCGGTQTRRVFCGIIVNNVYVESBK
CAGLTKYSEKRNCTVPKPCSLWYATQWSKCEGCGGTQTRRVFCGIIVNNVYVESBK
GLNVPDSQVQNCPTGCKNDEIITLDSHSTPIDEEDCDYEEYPTGLDFALAFSGRT
SPDLMSDTPSADIDSGTETGSDITDITVSDGIGIDDDGSGFSSFTDFTSDTD
DYSTDIDTSTVSSLSFSSVSDLKHTPRSENGTVKPKAAIATKDTTSTETSVTDL
TSETADTPTGLSTTSGDTSASDKSSTTDDKSTTDDFSSDSTITETYESGLTDSLTKE
SSLYATEDTSTSSLSFSSVSDLKHTPRSENGTVKPKAAIATKDTTSTETSVTDL
SSVADEETTKTITIEVSGTSEASINSKDTMTTLLSEDGTSVSEFDETTMSKVTPE
TSETSPTEGSDTKTMTSTVSEETSSVTEETTTTETTVENADTISSEVTSADKTTM
TMSSEKTEETEAITEITVKEVSESTETATSDKTISTLSEETKTSVAEESTTE
KVTEITVITMPEETGKTITSEIITKTVTEPTDVGSEAITSDKTISTVASETCK
YVSSEETVKTVAEASTEPSTEALISDKTKMSTLSPETKTSVSEIIVKTVTVEE
STEPSSEITSEKTSVISTSEETKTSVSEVTVKTVTDEATEITSTVSEIETKES
VSGSTEELSTOASSKIESPTSESTSTSESTSTSESTSTSESTSTSESTSTSEST
TETMLSTKEETVMSLTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST
STTVAPETNETGPASALPVEGLSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST
THRRCKPKRTCSALPVEGLSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST
GFPEGCPITLSDTTLFGCCLSDNKAAGNDGCGPPPPACKSTVFVGCADDETEA
RCDKKEGCPITLSTELSTSESTSESTSESTSESTSESTSESTSESTSESTSEST
CGFDETPATPGGEGGCGPCNTEFCGCPDGPVSPARGNDFEGCVTCNLSVGGCPDGE
TPAHGPDLLGCLLSAYGCCPNRPAVPHLEGGCCGYSGFCGCPDNATVARGPNFO
GGCGQVTEHGCDRHTEAAGPDYDGGCHTYFGCCPDGVTIAGKHQNGCDGRDQ
YGCQGDGTPATPEREGDCATSEYCCPDGLTEAKHKFYGCSDLFCVQAPKACNL
HORGTCNRYVYVYDLEYGCGRFWYGGCGNGKFKATKECEDVQVQAPKACNL
PKVKGACLYNIRWYDABEQCSQFIYGCGLGNANYSALQCOEQCPERSEDQCR
LPIDRSGSGNFRMGFNSERTRCQFMWGCENRFRSSEIACIQDPPGPKPLP
DICTLPAALDCADYRERYDTRKSCORFYGGCAGNNTAFAECGECSEBAKI
TITVRPTEAPLEMTFMEKDPGCTDITRNYDYKLGKCVTFYGGCGGNRNPPT
EYCYOYCTAQTICQLPMRSPCTESLMRNYDFSSDSCSOFTYGGCGDGNDRPETR
DCECVRSHNDLSDGSCRTHTVPTVAVDSIAEQLPTECHISPESEPCQSGQWYLDQ
RIGECVSHNDLSDGSCRTHTVPTVAVDSIAEQLPTECHISPESEPCQSGQWYLDQ
FEDKASRCPREFAYGCHGCHNRFSTIDECQVCRSEMDPKQVVEPQDCTSRVWY
YDNVDTCLQFIYGGCHGCHNRFSTIDECQVCRSEMDPKQVVEPQDCTSRVWY
TLPTTEAPTEAECPRESLPCGANITTYVYDOESRSCVQDQIGCRVANSFRTEE
CERRCGAFDRINVEARLDPCGRSSIPKRYEPASGSCLOYSGCGGAGNRFSTIE
ECEETCGNTGPDVVGPGPYEEELQLVCPGISRDPTPEGCTCACVAPVDCGPM
RVCEQMRCTYGMARTQDLNGCKSCLPDPCVRQRCPLGESCVAIAKBDPVTDAQF
YAVCKPVNEVLPCDRFNDCELSRLCEYVQVTRTVYDGCRCSCVDEIDCEBLKLEC
ERLKOYLVERQVADGCGCERSEHPCKNKECGPDRCVPTAYLDPTITQESRFSSEC
RPVNTKCYCPEQAATTEYPCNPECVDADRCGCGARGCGACACVAPDPTSAAPH
TTHAPDLPSPQVSOETPEVAEAGGKATLCIFPHNPPPKITWRRGEITIDGDRD
RYRLSDGSEIVSVYRNDSGVYICVAENDLGTAROFIYLVKVPVHTLAAASSLLT
AGSELVEVDYGPEDVITWKDRIIENDRIIRISGSSVSHLTVARVSLSDAGMYE
CHARNYSSSHYSYLOIAVPLTIPAKCTDSFFFAFNCLPIVRSFCQHKYYSKFCCKSC
LEAGOLSAHEVEMOQADQLVRK"
381..533
/note="Region: thrombospondin type I/properdin domain 1"
534..1187
/note="Region: PAL domain"
1227..1391
/note="Region: thrombospondin type I/properdin domain 2"
1401..1586
/note="Region: thrombospondin type I/properdin domain 3"
1593..1760
/note="Region: thrombospondin type I/properdin domain 4"
1773..1940
/note="Region: thrombospondin type I/properdin domain 5"
1950..2126
/note="Region: thrombospondin type I/properdin domain 6"
2136..2288
/note="Region: thrombospondin type I/properdin domain 7"
misc_feature
381..533
misc_feature
534..1187
misc_feature
1227..1391
misc_feature
1401..1586
misc_feature
1593..1760
misc_feature
1773..1940
misc_feature
1950..2126
misc_feature
2136..2288
```

```
misc_feature 2289..4886
/notes="Region: TES domain"
misc_feature 4887..4958
/notes="Region: lagrin domain 1"
misc_feature 4974..5045
/notes="Region: lagrin domain 2"
misc_feature 5058..5132
/notes="Region: lagrin domain 3"
misc_feature 5211..5222
/notes="Region: lagrin domain 4"
misc_feature 5352..5423
/notes="Region: lagrin domain 5"
misc_feature 5427..5498
/notes="Region: lagrin domain 6"
misc_feature 5508..5579
/notes="Region: lagrin domain 7"
misc_feature 5580..5651
/notes="Region: lagrin domain 8"
misc_feature 5655..5726
/notes="Region: lagrin domain 9"
misc_feature 5730..5801
/notes="Region: lagrin domain 10"
misc_feature 5805..5876
/notes="Region: lagrin domain 11"
misc_feature 5880..5951
/notes="Region: lagrin domain 12"
misc_feature 5955..6026
/notes="Region: lagrin domain 13"
misc_feature 6057..6209
/notes="Region: Kunitz proteinase inhibitor domain 1"
misc_feature 6234..6386
/notes="Region: Kunitz proteinase inhibitor domain 2"
misc_feature 6411..6563
/notes="Region: Kunitz proteinase inhibitor domain 3"
misc_feature 6597..6749
/notes="Region: Kunitz proteinase inhibitor domain 4"
misc_feature 6810..6962
/notes="Region: Kunitz proteinase inhibitor domain 5"
misc_feature 6981..7133
/notes="Region: Kunitz proteinase inhibitor domain 6"
misc_feature 7224..7385
/notes="Region: Kunitz proteinase inhibitor domain 7"
misc_feature 7410..7562
/notes="Region: Kunitz proteinase inhibitor domain 8"
misc_feature 7581..7733
/notes="Region: Kunitz proteinase inhibitor domain 9"
misc_feature 7848..7997
/notes="Region: Kunitz proteinase inhibitor domain 10"
misc_feature 8022..8174
/notes="Region: Kunitz proteinase inhibitor domain 11"
misc_feature 8208..8516
/notes="Region: antistatin-like domain 1"
misc_feature 8544..8852
/notes="Region: antistatin-like domain 2"
misc_feature 8868..8999
/notes="Region: WAP/4 disulfide domain"
misc_feature 9105..9281
/notes="Region: immunoglobulin domain 1"
misc_feature 9378..9557
/notes="Region: immunoglobulin domain 2"
misc_feature 9627..9722
/notes="Region: PLAC domain"
BASE COUNT 3109 a 2750 c 2738 g 2411 t
ORIGIN

alignment_scores:
  Quality: 141.00 Length: 118
  Ratio: 2.136 Gaps: 6
Percent Similarity: 55.932 Percent Identity: 30.508

alignment_block:
US-09-215-435-177 x AF078161
```

```
Align seg 1/1 to: AF078161 from: 1 to: 11008

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhe..... 12
:::|||||:::|||||::: |||||::: |||||:::
6249 GTCAGGAGGCGCTGTTGGGCTACAATATTCGCTGCTACTACGACGCCGA 6298

13 .....PheIlePhe.....CysSerLeuAsnThrL 21
|||||::: |||||::: |||||:::
6299 GCAGGAGCAATGCAGTCAGTTTCATCTATGGGGTTGT ..... 6335

21 euLeuLeuGlyValAlaAsnLysIleAlaGlu...LysIleCysGlyAsp 36
|||||:::|||||::: |||||::: |||||::: |||||:::
6336 .....CTTGGCAATGCAACAACACTACGCTCCTCACTTTCGCAAGAG 6380

37 .....LeuLysAspProCysLysLysLeuAspMetAsnPh 47
:::||||| |||||::: |||||::: |||||:::
6381 CAATGCTCAACCTGAAGGAGTGAAGATCAATGCCGCTGCCAATTGACGC 6430

47 eGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerL 64
|||||::: |||||::: |||||::: |||||::: |||||:::
6431 AGGTCCTTTCCGGCAACTTTGGCGCGCTTCAACAGCGAAACGC 6480

64 ysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsn 80
:::|||||::: |||||::: |||||::: |||||::: |||||:::
6481 GACGCTGCGAGCAGTTCATGTGGGCGGTTGTGAGGGCAACACCAATAGG 6530

81 PheLysLeuLysIleGluArgGluValAlaLysValAlaLysTyrLyspr 97
|||::: |||||::: |||||::: |||||::: |||||:::
6531 TTC.....AGCTCTGAGATCGCTTGCATCCAGCGTTGTGACCC 6568

97 oPro 98
|||||
6569 ACCA 6572

seq_name: gb_pat2:E12656

seq_documentation_block:
LOCUS E12656 555 bp DNA PAT 24-JUN-1998
DEFINITION Synthetic DNA encoding urine trypsin inhibitor linked with SUC2
signal.
ACCESSION E12656
VERSION E12656.1 GI:3251488
KEYWORDS JP 1997047295-A/1.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 555)
AUTHORS Ooya,T., Morita,M., Tanabe,T., Miura,M. and Kobayashi,K.
TITLE PRODUCTION OF PROTEIN
JOURNAL Patent: JP 1997047295-A 1 18-FEB-1997;
GREEN CROSS CORP:THE
COMMENT
OS None
OC Artificial sequences.
PN JP 1997047295-A/1
PD 18-FEB-1997
PF 08-AUG-1995 JP 1995224690
PI Ooya TOMOSUKE, MORITA MASANORI, TANABE TOSHIZUMI, MIURA
MASAMI, PI KOBAYASHI KAORU
PC C12P21/02.C12NI/19,C12NI5/09,(C12P21/02,C12R1:84),(C12NI/19,
C12R1:84);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..555
FT /organism='Artificial sequences' FT
FT sig_peptide 1..60
FT /product='SUC2 signal peptide' FT
FT mat_peptide 61..504
FT /product='urine trypsin inhibitor'.
```

```
FEATURES
  source      Location/Qualifiers
              1..555
              /organism="unidentified"
              /db_xref="taxon:32644"
BASE COUNT   121 a 143 c 164 g 127 t
ORIGIN
alignment_scores:
  Quality: 135.50      Length: 99
  Ratio: 2.710        Gaps: 3
  Percent Similarity: 50.505  Percent Identity: 36.364
alignment_block:
  US-09-215-435-177 x E12656
  Align seg 1/1 to: E12656 from: 1 to: 555
9 LeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuGlyG1 25
|||||:  |||:|||||
7 CTTTGCACAGCTTCCTTTC.....CTTTGGCTGGTTT 41
25 yValAsnLysIleAlaGluLysIleCys..... 34
42 TGCAGCCAAATATCTGCAGCTGTCTACCCCAAGAGAGGATCAG 91
35 .....GlyAspLeu.....LysAspProCysLys 42
92 GGGGTGGGCACTGGTAAGTCAAGTACCACAGAAAGAGATTCCTGCCAG 141
43 LeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheTy 59
|||||:  |||:|||||
142 CTGGGCTACTCGCGCGTCCCTGCATCGGGAATGACCAGCAGGTATTCTTA 191
59 rAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsnG 76
|||||:  |||:|||||
192 TAATGGTACATCCATGGCCTGTGAGACTTTCCAGTACGGGGCTGCATGG 241
76 yValAsnLeuAsnPhelLysLeuLysIleGluArgGluValAlaCys 91
|||||:  |||:|||||
242 GCACGGTAACAACCTTCGTACACAGAAAGAGGTCTGTGCAGACCTGC 288
seq_name: gb_om:BTU35642
seq_documentation_block:
LOCUS      BTU35642      1198 bp      mRNA      MAM      05-SEP-1996
DEFINITION Bos taurus alphas 1-microglobulin/bikunin mRNA, complete cds.
ACCESSION  U35642
VERSION    U35642.1  GI:1016297
KEYWORDS
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE
  1 (bases 1 to 1198)
  Lindqvist, A. and Akerstrom, B.
  Bovine alpha 1-microglobulin/bikunin. Isolation and
  characterization of liver cDNA and urinary alpha 1-microglobulin
  Biochim. Biophys. Acta, Gene Struct. Expr. 1306 (1), 98-106 (1996)
  96201710
REFERENCE
  2 (bases 1 to 1198)
  Lindqvist, A.
  Direct Submission
  Submitted (08-SEP-1995) Annika Lindqvist, Molecular Signalling,
  Lund University, P.O. Box 94, Lund, 221 00, Sweden
  Location/Qualifiers
  1..1198
  /organism="Bos taurus"
  /db_xref="taxon:9913"
  /tissue_type="liver"
  104..1162
  /note="precursor protein for alphas 1-microglobulin and
  CDS
  bikunin"
  /codon_start=1
  /product="alpha 1-microglobulin/bikunin"
  /protein_id="AAB07599.1"
  /db_xref="GI:1016298"
  /translation="MRSISGLLLLTACLAVNASSVPTLPDDIQVOENFDLSRIYGKW
  FNVAVGTCPLKREKEMTSTVYVLIAGPSTKEISVTNTHRRKGVCESISGTVEKTS
  ADGKFLYHAKAKWNTMESVYVHTNDEVALFTKLKSRHGGTTIVKLYGREQLRES
  LLEEFREVALGVGIPEDALFTMPDRGECVPGQDPVPTPLSRARRAVLTQEESSGAG
  QVPTNFSKKACDQDLSQSGPCGLGFLKRYFYNGTSMACETFLYGCNGNNGNF,LSKEKE
  CLQTCRTVEACNLPIVQGPCRSYIQLWAFDAVKGKCVRFSGYGGCKNGNKFYSEKECK
  EYCGIPGEADELLRFN"
BASE COUNT   284 a 308 c 354 g 252 t
ORIGIN
alignment_scores:
  Quality: 134.00      Length: 43
  Ratio: 4.467        Gaps: 0
  Percent Similarity: 69.767  Percent Identity: 55.814
alignment_block:
  US-09-215-435-177 x BTU35642
  Align seg 1/1 to: BTU35642 from: 1 to: 1198
39 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 55
|||||:  |||:|||||
788 GATTCCTGCCAGCTGCACACAAAGGCGCTTGCCTGGGCTGTTCACAA 837
55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhes 72
|||||:  |||:|||||
838 GAGGTATTTCTATAACGGTATACATCCATCGCCTGTGAGACCTCTTATATG 887
72 erGlyCysAsnGlyAsnLeuAsnAsnPh 81
:::|||||:  |||:|||||
888 GCGGCTGCATGGGAATGGCAACAACCTTC 916
seq_name: gb_in3:CELC37C3
seq_documentation_block:
LOCUS      CELC37C3      42521 bp      DNA      INV      08-JUL-1998
DEFINITION Caenorhabditis elegans cosmid C37C3.
ACCESSION  U64857
VERSION    U64857.1  GI:1465829
KEYWORDS
SOURCE     Caenorhabditis elegans.
            Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
            1 (bases 1 to 42521)
            Wilton, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
            Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
            Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
            Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
            Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
            Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
            O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Ropstra, A.,
            Saunders, D., Shownkeen, R., Smaildon, N., Smith, A., Sonnenhammer, E.,
            Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
            Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
            Wilkinson-Sproat, J. and Wohlman, P.
            2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
            elegans
            Nature 368 (6466), 32-38 (1994)
            94150718
            2 (bases 1 to 42521)
            Geisel, C. and Bradshaw, H.
            The sequence of C. elegans cosmid C37C3
            Unpublished (1998)
            3 (bases 1 to 42521)
            Waterston, R.
            Direct Submission
            TITLE
            JOURNAL
            MEDLINE
            REFERENCE
            AUTHORS
            TITLE
            JOURNAL
            MEDLINE
            REFERENCE
            AUTHORS
            TITLE
```



```

JOURNAL Submitted (22-JUL-1996)
REFERENCE 4 (bases 1 to 42521)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1998) Department of Genetics, Washington
COMMENT University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
Submitted by: Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rw@ematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is 2K742, 200 bp overlap; 3' cosmid is F40A3, 200 bp
overlap. Actual start of this cosmid is at base position 1 of
CELC37C3; actual end is at 5024 of CELF40A3

NOTES:

Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in preparation).
Location/Qualifiers
1..42521
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/clones="C37C3"
/chromosome="V"
513..15440
/feature="C37C3.6b"
/feature="C37C3.6a"
join(513..608,664..878,1380..1535,1583..1734,1785..2541,
2666..2897,2976..3097,3149..4733,4801..5584,5659..5871,
6194..6558)
/feature="C37C3.6a"
/feature="similar to the DPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor;
coded for by C. elegans cDNA yk84d8.5; coded for by C.
elegans cDNA yk117d1.5; coded for by C. elegans cDNA
yk86e5.3; coded for by C. elegans cDNA yk132d4.3; coded
for by C. elegans cDNA yk123e8.3; coded for by C. elegans
cDNA yk88e5.5; coded for by C. elegans cDNA yk22b2.5;
coded for by C. elegans cDNA yk132d4.5; coded for by C.
elegans cDNA yk123e8.5; coded for by C. elegans cDNA
CEESD32F; coded for by C. elegans cDNA CEES31FC; coded
for by C. elegans cDNA CEES167F; coded for by C. elegans
cDNA CEES173F; coded for by C. elegans cDNA CEES33F;
coded for by C. elegans cDNA CEES94F; coded for by C.
elegans cDNA CESAC89F; coded for by C. elegans cDNA
yk10f9.3; coded for by C. elegans cDNA yk10f9.5; coded for
by C. elegans cDNA yk22b2.3; coded for by C. elegans cDNA
yk3e8.3; coded for by C. elegans cDNA yk3e8.5; coded for
by C. elegans cDNA yk41h8.5; coded for by C. elegans cDNA
yk5e8.5; coded for by C. elegans cDNA yk148d2.5; coded for
by C. elegans cDNA yk152f10.5"

/feature="similar to the BPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor
(TFPI); coded for by C. elegans cDNA yk257g8.5; coded for
by C. elegans cDNA yk257g8.3; coded for by C. elegans cDNA
yk185b12.3"
/codon_start=1
/protein_id="AAC25868.1"
/db_xref="GI:3294502"
/translation="MRLLLSAALLLCVPTWAFSLSSFFSGSDVAQKPYLHPNSPPER
DPASSRMKRAYQYVVDGVSVDKSGQKETGNWGPWPVNECSRSGCGGVQLEKQ
CSGDTGASRYIISCNLACSESGDFRAEQSKENDEALDGNHKTGYKGNKCELV
CPKESGNFYKWDKVDVDTKDSKNDICVDGECIPVGCDCGLSSLKFKCGKCDG
DGSCTKIIEGRFDERNLSPGYHDIILKPEGATNIQIARKSTNNLALNKGSDHFLN
GNGLIOVEKEVEVGSTIFVYDDAEPETLSAOGPLSELTVALLRKGRSDPTAIKVEFS
IPLLEEDYMYKFDNWTPCSVSCGKGVQTRNLICIDGKNKGRVEDDLCEENNAKPEF
EKSCETVDCAEWETGWSKSSCTGGGQGYRVYVCHQFANRRVTVEDGCTVTER
PVKOTCNRFACPEWQAGPWSACEKGDFAFYSVTCRSKEEGEEGLAADAACPAD
EQEFDETRTCNLGPEGLTFVTGEMNLCTRCNDTEETREVTCKDSQGRAYPLEKCLV
DNSTEIPDTRSCATQPCYEYVTSWSEKCTTECGHGHKTRRVICAHONGGLEVPE
EGHQAEPKGTNCTNEKCTGTWYTSWSECTAECGGSDQAVAVCLNVDKKKPVDE
WCDRAVPSEKQDCNVDCTPCVDSEFGCCPDNSTFATGEFNFSGCSNSETFECCAD
NVTATGPNSGKEEFVESPUNLEADVANADAEASGDAPELCSVTNNGEAVSFTLGGC
IAPITALLGDELIGNDTASNETIHCSKTFEGCCPDWYTAASGKNEGCPSFTLGGC
NETQFGCHDDVTIARGANLGGCEPSCAASLYGCKDRKTIAGFPHYSGERSFPFC
ELSDFGCCPDGETAALGKNGTGCENCLTTKFGCCPDGKTKAKSHNEGCGCEFAQY
CCPDGKSVAKGAGFYGCPESCAOSQFCCPDGKTRARGENKGCPCOYTRYGCCPDGE
TTALGPRNDGCDNCRYAKHGCCPDGETKALGPDGAGCPTTTPFLMGSTVAPHKIAA
CNQTESGTCVAGYKLAHWDTTEGRCNQFWYGGCGNDNNAFASQDMCTTICVEPPG
KGRYLPDPRVGLPCRDQLPRYYDDHSHKHCVAFWRGCLQGNANFNNSFECSEFKD
VGPYDAPTAAPPPQDQNAQYLPTEVQIEIQSAEQPQPOQOQOQOQOQPOQ
PROSMEDICRSRODAGPCETYSDOWFYNAFSECECTFYTGCGGNLNRFRSKDCEOR
CFVHGAGPSAARQEQAPQAQPAQPSNIVSPQOQASPVVVPVPSNKSORDACHLN
VDQGRCKGAFDSWYEVATGSCVTFKTYGCGGNANRFASKQCESLCVKPASEASAG
IDGAAGINSVCDEAKDTPCFNFTWYKYNKADGTCNRFHYGGCGGTNRFRATIEBOCK
AACQNHKDACQLPRVQGPCSGKSHYYNTASHQCTETFTYGGCGLGNTNRFRATIEBOCK
RCPSEK"
join(513..608,664..878,1380..1535,1583..1734,1785..2541,
2666..2897,2976..3097,3149..4733,4801..5584,5659..5871,
6194..6547,7038..7134,7881..8009,8060..8410,8493..8571,
13650..13741,14139..14727,14940..15440)
/feature="C37C3.6b"
/feature="similar to the BPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor
(TFPI); coded for by C. elegans cDNA yk257g8.5; coded for
by C. elegans cDNA yk257g8.3; coded for by C. elegans cDNA
yk185b12.3"
/codon_start=1
/protein_id="AAC25868.1"
/db_xref="GI:3294502"
/translation="MRLLLSAALLLCVPTWAFSLSSFFSGSDVAQKPYLHPNSPPER
DPASSRMKRAYQYVVDGVSVDKSGQKETGNWGPWPVNECSRSGCGGVQLEKQ
CSGDTGASRYIISCNLACSESGDFRAEQSKENDEALDGNHKTGYKGNKCELV
CPKESGNFYKWDKVDVDTKDSKNDICVDGECIPVGCDCGLSSLKFKCGKCDG
DGSCTKIIEGRFDERNLSPGYHDIILKPEGATNIQIARKSTNNLALNKGSDHFLN
GNGLIOVEKEVEVGSTIFVYDDAEPETLSAOGPLSELTVALLRKGRSDPTAIKVEFS
IPLLEEDYMYKFDNWTPCSVSCGKGVQTRNLICIDGKNKGRVEDDLCEENNAKPEF
EKSCETVDCAEWETGWSKSSCTGGGQGYRVYVCHQFANRRVTVEDGCTVTER
PVKOTCNRFACPEWQAGPWSACEKGDFAFYSVTCRSKEEGEEGLAADAACPAD
EQEFDETRTCNLGPEGLTFVTGEMNLCTRCNDTEETREVTCKDSQGRAYPLEKCLV
DNSTEIPDTRSCATQPCYEYVTSWSEKCTTECGHGHKTRRVICAHONGGLEVPE
EGHQAEPKGTNCTNEKCTGTWYTSWSECTAECGGSDQAVAVCLNVDKKKPVDE
WCDRAVPSEKQDCNVDCTPCVDSEFGCCPDNSTFATGEFNFSGCSNSETFECCAD
NVTATGPNSGKEEFVESPUNLEADVANADAEASGDAPELCSVTNNGEAVSFTLGGC
IAPITALLGDELIGNDTASNETIHCSKTFEGCCPDWYTAASGKNEGCPSFTLGGC
NETQFGCHDDVTIARGANLGGCEPSCAASLYGCKDRKTIAGFPHYSGERSFPFC
ELSDFGCCPDGETAALGKNGTGCENCLTTKFGCCPDGKTKAKSHNEGCGCEFAQY
CCPDGKSVAKGAGFYGCPESCAOSQFCCPDGKTRARGENKGCPCOYTRYGCCPDGE
TTALGPRNDGCDNCRYAKHGCCPDGETKALGPDGAGCPTTTPFLMGSTVAPHKIAA
CNQTESGTCVAGYKLAHWDTTEGRCNQFWYGGCGNDNNAFASQDMCTTICVEPPG
KGRYLPDPRVGLPCRDQLPRYYDDHSHKHCVAFWRGCLQGNANFNNSFECSEFKD
VGPYDAPTAAPPPQDQNAQYLPTEVQIEIQSAEQPQPOQOQOQOQOQPOQ
PROSMEDICRSRODAGPCETYSDOWFYNAFSECECTFYTGCGGNLNRFRSKDCEOR
CFVHGAGPSAARQEQAPQAQPAQPSNIVSPQOQASPVVVPVPSNKSORDACHLN
VDQGRCKGAFDSWYEVATGSCVTFKTYGCGGNANRFASKQCESLCVKPASEASAG
IDGAAGINSVCDEAKDTPCFNFTWYKYNKADGTCNRFHYGGCGGTNRFRATIEBOCK
AACQNHKDACQLPRVQGPCSGKSHYYNTASHQCTETFTYGGCGLGNTNRFRATIEBOCK
RCPSEK"

```

```

* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 134095: contig of 134095 bp in length.
*
FEATURES             source
1..134095            /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
                     /clone="Y97E10y"
BASE COUNT          43230 a 24499 c 23702 g 42664 t
ORIGIN

alignment_scores:
  Quality: 133.00      Length: 62
  Ratio: 3.500        Gaps: 0
  Percent Similarity: 61.290 Percent Identity: 40.323

alignment_block:
US-09-215-435-177 x AC006915
..
Align seg 1/1 to: AC006915 from: 1 to: 134095
34 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCy 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6370 TGTCAAAATCATAGGATGCTGTCAACTTCCAAAGGTTCCAGGACCATG 54719

50 styrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6420 CTCGGAAGCATCTCTATTATTACTACAACTGCCAGTCATCATGCG 54769

67 luThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54770 AGAGTTCTACTATGTTGGCTGCTCGGAATACTACAGATTCGCTACC 54819

84 LysileGluArgGluValAlaCysValAlaLysTyr 95
|||||:|||||:|||||:|||||:|||||:|||||:
54820 ATTGAGGAGTGTCAAGCGAGATGCCCGAGTAAGTTC 54855

seq_name: gb_htg1.AC006915
seq_documentation_block: 46847 bp DNA HTG 03-JAN-2000
LOCUS AC006915 Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
ACCESSION AC006915
VERSION AC006915.1 GI:6664904
KEYWORDS HTG: HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 46847)
AUTHORS Adams, M. and Venter, J.C.
JOURNAL Direct Submission
Rockville, MD, USA
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10211667 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
FEATURES             source
1..46847            /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT          13884 a 9574 c 9324 g 14065 t

```

```

TSCVAKIERVESVSASKSLEEVRLTDPRDSHFYHDPVDOIEAEYVIVDTGA
LPCLMPLQGRSCYDNLIRPFDSEKSCQVTFNMGCGNPNANHTSOETCEACGKW
RNVAVCLEPAEHGDCOLAIPIRYHDPKTSQCMWMTGCGNGNAPSSKADCESLRV
ETLWNTDFTCLERSAGPCTDSISMWYFDTHLCKPFTYGGCRGNRRVSEKQCO
OSCRPGDTSEDICTLRPPEPCRLGLEKIFYDPIQSCMHFYGGCGEGNANRDSSEL
DCPRGSSVKVEASESERVGLTASSTPIYIVNKTATFVGNTRFRICNSGVLPITW
YKNGLLQFOSGRITERENDTLEIVALTADAGVCTIAGQDSTMSEGVVYIKRLPGH
RTTSRMLPTPSKRNPSLGTPTSPSTVSTPRIVTPGSPSDARVSRTNSCNDVG
NASTCDLIVKNGLCGRKRYGTCFCHTCIRVHNEKF"
complement(8582..12734)
/gene="C37C3.7"
complement(join(8582..8744,8795..8868,8914..9194,
9378..9558,9607..9722,10028..10144,10473..10730,
10960..11122,11912..12280,12651..12734))
/gene="C37C3.7"
/notes="contains similarity to EGF-like domains
(PS:PS00022)"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC25862.1"
/db_xref="GI:1465832"
/translation="MTWSEKQEYCRIGSNDLNTCTCVGKSGNCFWCGGKTRKCMPPD
..
alignment_scores:
  Quality: 133.00      Length: 62
  Ratio: 3.500        Gaps: 0
  Percent Similarity: 61.290 Percent Identity: 40.323

alignment_block:
US-09-215-435-177 x CELC37C3
..
Align seg 1/1 to: CELC37C3 from: 1 to: 42521
34 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCy 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6370 TGTCAAAATCATAGGATGCTGTCAACTTCCAAAGGTTCCAGGACCATG 6419

50 styrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6420 CTCGGAAGCATCTCTATTATTACTACAACTGCCAGTCATCATGCG 6469

67 luThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6470 AGAGTTCTACTATGTTGGCTGCTCGGAATACTACAGATTCGCTACC 6519

84 LysileGluArgGluValAlaCysValAlaLysTyr 95
|||||:|||||:|||||:|||||:|||||:|||||:
6520 ATTGAGGAGTGTCAAGCGAGATGCCCGAGTAAGTTC 6555

seq_name: gb_htg1.AC006915
seq_documentation_block:
LOCUS AC006915 134095 bp DNA HTG 24-FEB-1999
DEFINITION Caenorhabditis elegans clone Y97E10y, *** SEQUENCING IN PROGRESS
ACCESSION AC006915
VERSION AC006915.1 GI:4263433
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodrilidae; Caenorhabditis.
1 (bases 1 to 134095)
AUTHORS Waterston, R.H.
JOURNAL Unpublished
Rockville, MD, USA
Submitted (24-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
* NOTE: This is a 'working draft' sequence. It currently

```


AC005149 98942 bp DNA INV 20-JUN-1998
Drosophila melanogaster DNA sequence (Pls DS03364 (D237) and
DS02911 (D145)), complete sequence.
AC005149 AC004339 AC004378 AC003204 AC003205 AC002788 AC002789
AC002790 AC002791 AC002792 AC002793 AC002794 AC002795 AC002796
AC003206 AC002797 AC002798 AC002799 AC003207 AC002800 AC002801
AC002802 AC003208
AC005149.1 GI:3242679
HTG.
Drosophila melanogaster (Subclones in tet from P1 clones DS03364
(D237) and DS02911 (D145)) DNA.
ORGANISM
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 98942)
Celniker,S.E., Aghavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Humasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
Lomutan,M.A., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M.,
Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,D., Snir,E.,
Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.L.
and Kimmel,B.
Direct Submission
Submitted (20-JUN-1998) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site
(http://fruitfly.berkeley.edu/sequence-archive.html) or send email
to drosophila@hgsc.lbl.gov.
Library locations: 4.36, 31.31.
Location/Qualifiers
1 . 98942
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
/map="2A1-24C5"
/clone="Pls DS03364 (D237) and DS02911 (D145)"
/note="DS03364 (D237) extends from bp 1 to bp 79652 and
DS02911 (D145) extends from bp 15716 to bp 98943"
BASE COUNT 25273 a 20900 c 20166 g 28603 t
ORIGIN

alignment_scores:
Quality: 132.50 Length: 101
Ratio: 2.366 Gaps: 4
Percent Similarity: 55.446 Percent Identity: 31.683

alignment_block:
US-09-215-435-177 x AC005149/rev ..
Align seg 1/1 to reverse of: AC005149 from: 1 to: 98942
12 prepHeliePhCySerLeuAsnThrLeuLeuGlyVal..... 26
||||| :|||:

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

27AsnLysIleAlaGluLys.....IleCysG 35
2549 AGCTCCCTCGATTCGTGGTGAAGCGGAGGAGCGCGCGCGGCCA 2500
35 lyAspLeuLysasp.....ProCysLysLeu 43
2499 AACCCCTGCCTGACTGCAGTGAGCTCACTGTTCCGGAGGACTGTTCATCAG 2450
44 AspMetAsnPheGlySerCystyrGluValHisPheArgTyrPheTyrAs 60
2449 CCCAAGGAACCCTGCTGCTTCGCTCTGTTCTATCGCTACGCTACAA 2400
60 nArGThrSerLysArGcysGluThrPheValPheSerGlyCysAsnGlyA 77
2399 CGTGGATACACAATCCCTCGGAGGAGTTCGTTTACGGTGGATGTCGGCGCA 2350
77 snLeuAsnAsnPheLysLeuLysileGluArgGluValAlaCysValAla 93
2349 ACAAGAACAACCTTCGATCCAGGACGAGTGGCAACAGCATGTTTGGTA 2300
94 Lys 94
2299 AAG 2297
seq_name: gb_inl:AE003579
seq_documentation_block:
LOCUS AE003579 304383 bp DNA INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386046 section 12
of 16, complete sequence.
ACCESSION AE003579 AE002638
VERSION AE003579.2 GI:10727349
KEYWORDS HTG.
SOURCE fruit fly,
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 304383)
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.C., Wortman,J.R., Yeager,M.D., Zhang,Q., Chen,L.X., B.D.,
Brandon,R.C., Rogers,Y.H., Blazer,R.G., Chame,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
Miklos,G.L., Abell,J.F., Agbayani,A., An,H.J.,
Andrews-Pfannkuch,C., Baldwin,D., Balow,R.M., Basu,A.,
Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
Benos,P.V., Bernan,B.P., Brokstein,P., Brotter,P., Burtis,K.C.,
Botchan,M.R., Bouck,J., Cadieu,E., Center,A., Chandra,I.,
Busam,D.A., Butler,H., Dahlke,C., Davenport,L.B., Davies,P., de
Cherry,J.M., Cawley,S., Dahlke,C., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
Pablos,B., Delcher,A., Deng,Z., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
Dudson,K., Doup,L.E., Evansgelist,C., Ferraz,C., Dunn,P., Durbin,K.J., Evangelista,C., Gabrielian,A.E., Garg,N.S.,
Fleischmann,W., Fosler,K., Glaser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
Hernandez,J.R., Houck,J., Hostin,D., Huo,K., Howland,T.J.,
Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Milshina,N.V., Mobarry,C.,
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Murphy,L.,
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.L., Nelson,D.R., Palazzolo,M., Pittman,K.A., Nixon,K.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Klamos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R.,
Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A.,


```

Quality: 131.00      Length: 85
Ratio: 2.911         Gaps: 2
Percent Similarity: 52.941      Percent Identity: 37.647

alignment_block:
..
US-09-215-435-177 x I63559
from: 1 to: 1455

```

```

10 TTATTATTCGAATCTTCTTAGTGTTGCGCCGGCCTGCAGTGTGTTGAAAGTTGTTTAG
    :      :      :      :      :      :      :      :      :      :      :
38 sAsPrOcYsLysLeuAspMetasnPheGlySerCystyrGluvalHlsp 55
    |||:::|||   ::|||   |||   |||   |||   |||   |||   |||   |||
60 AGACTCTGCACGTGGGTACTCGGCCGGTCCCTGTCATCGAATGACCA 109
    :      :      :      :      :      :      :      :      :      :
55 heArgTyrrPheTyrsnArgrThrSerLysArgCysgluThrrPheValPhe 71
    |||:::|||   |||   |||   |||   |||   |||   |||   |||   |||
110 GCAGGTATCTATAATGTGCATCCATCGCTGCGTGTGAGACTTTCAGTAC 159
    :      :      :      :      :      :      :      :      :      :
72 SerGlyCyasaSnGlyAsnLeuAsnAsnPheLysLeuLysalleclu.... 86
    |||:::|||   |||   |||   |||   |||   |||   |||   |||   |||
160 GCGCGTCGTCATGGCGCAACGGTAACAACCTTCGCACAGAAAAGCAGTGTCT 209
    :      :      :      :      :      :      :      :      :      :
87 .....ArgGluVal.....AlaCysVala 93
    |||   |||   |||   |||   |||   |||   |||   |||   |||
210 GCAGAGCTGCGCGAACACTGTGGCGCGCGCTGAAACTGTTGAAAGTTGTTTAG 259
    :      :      :      :      :      :      :      :      :      :
93 laLys 94
    |||   |||   |||   |||   |||   |||   |||   |||   |||
260 CAAAA 264

```

```

seq_name: gb_om:MEELACP

seq_documentation_block:
LOCUS           419 bp          mRNA          MAM          15-APR-1998
DEFINITION      Macropus eugenii mRNA for early lactation protein (ELP).
ACCESSION       AJ000490
VERSION         AJ000490.1 GI:3059123
KEYWORDS        early lactation protein; ELP gene.
SOURCE          tamar wallaby.
ORGANISM        Macropus eugenii
                 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                 Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

REFERENCE
AUTHORS         Simpson,K.J.
TITLE           Direct Submission
JOURNAL         Submitted (17-JUL-1997) Simpson K.J., Division of Molecular Biology
                 and Genetics, Victorian Institute of Animal Science, 475 Mickleham
                 Rd, Attwood, Victoria, 3049, AUSTRALIA
REMARK          Revised by [3]

REFERENCE
AUTHORS         Simpson,K.J., Shaw,D. and Nicholas,K.R.
TITLE           Asynchronous expression of a putative protease inhibitor gene in
                 the mammary gland during lactation in the Tamar wallaby, Macropus
                 eugenii
JOURNAL         Unpublished
REFERENCE
AUTHORS         Simpson,K.J.
TITLE           Direct Submission
JOURNAL         Submitted (14-APR-1998) Simpson K.J., Division of Molecular Biology
                 and Genetics, Victorian Institute of Animal Science, 475 Mickleham
                 Rd, Attwood, Victoria, 3049, AUSTRALIA
COMMENT          On Apr 18, 1998 this sequence version replaced gi:3046757.
FEATURES
source          1..419
                 /organism="Macropus eugenii"
                 /db_xref="taxon:9315"
gene            1..252
                 /gene="ELP"
CDS             <1..252
                 /gene="ELP"

```

```

/codon_start=1
/evidence-experimental
/product="early lactation protein"
/protein_id="CAA04128.1"
/db_xref="GI:3046758"
/db_xref="SPTREMBL:O62845"
/translation="EXCLDOIQVNSLENLSLVPSLCLLPVPRGNCSSQLLHIFYNT
SRCTFFIYSGNGNRNNNSEYCLKTRNNRNNN"
source
1..255
/organism="Macropus eugenii"
/db_xref="taxon:9315"
/tissue_type="lactating mammary gland"
polyA_signal
polyA_site
386
BASE COUNT 139 a 98 c 72 g 110 t
ORIGIN

alignment_scores:
  Quality: 130.00      Length: 41
  Ratio: 4.333        Gaps: 0
Percent Similarity: 73.171 Percent Identity: 56.098

alignment_block:
US-09-215-435-177 x MEELACP ..
Align seg 1/1 to: MEELACP from: 1 to: 419

41 CysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTy 57
|||||
|||||
67 TGTCGTCTCCACCTGTGAGGGTAACCTGCAGAGTTCTCAAAATCTTCACTA 116
|||||
|||||
57 rPheTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyC 74
|||||
|||||
117 TTTCTACACACACCTCTCGAAGCTGTGAGACTTTCATCTACAGTGGCT 166
|||||
|||||
74 ysAsnGlyAsnLeuAsnAsnPhe 81
|||||
167 GCAATGGCAATAGGAACAATTTT 189

seq_name: gb_pr8:HSALMICR

seq_documentation_block:
LOCUS HSALMICR 1221 bp mRNA PRI
DEFINITION Human mRNA for alpha-1-microglobulin and HI-30.
ACCESSION X04494
VERSION X04494.1 GI:24478
KEYWORDS alpha-1-microglobulin; glycoprotein; proteinase inhibitor; serum
protein; signal peptide.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Kaumeyer,J.F., Polazzi,J.O. and Kotick,M.P.
TITLE The mRNA for a proteinase inhibitor related to the HI-30 domain of
inter-alpha-trypsin inhibitor also encodes alpha-1-microglobulin
(protein HC)
JOURNAL Nucleic Acids Res. 14 (20), 7839-7850 (1986)
MEDLINE 87040757
COMMENT Data kindly reviewed (11-MAY-1987) by Polazzi J.O.
FEATURES
source
1..1221
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
46..1104
/note="precursor polypeptide"
/codon_start=1
/protein_id="CAA28182.1"
/db_xref="GI:24479"
/translation="MRSGLGALLLLSACLAVSAGPVPPTPDNTQVQENFNSIRTYGK

```

YNLAIAGTSPWLKKINDRMVTSTVLVLEGATEAEIISMTSTRWRKGVCEETSGAYEKRTI TDGKFLYHRSKNWITMSXVVHTNYDEYAIETLTKFSRHGGPTITAKLYGRAPLRET LIQDFRVAQGVIGPDSISFTWADRGECVPGEPEPILIPRVRRVAVLPQEEGSGG OLVTEYTKKEDSCOLGYSAGPCMGMTSRVFNCTSMACETFOYGGCMGNFNFTKE CLQTCRTVAACNLPIVGRGPCRAFIQLWAFDAVKGKCVLFPPYGGCGNGNKFYSERKCR EYCGVPGDGDDELLRFSN"									
sig_peptide	46..102	/product="pot. signal peptide (AA -19 to -1)"							
mat_peptide	103..1101	/product="alpha-1-microglobulin-HI-30 protein (AA 1-333)"							
mat_peptide	103..651	/product="alpha-1-microglobulin"							
mat_peptide	661..1095	/product="HI-30"							
misc_feature	724..891	/note="domain I"							
misc_feature	892..1095	/note="domain II"							
misc_feature	1188..1192	/note="put. polyA signal"							
BASE COUNT	296 a	322 c	363 g	240 t					
ORIGIN									
alignment_scores:	Quality: 129.50	Length: 70							
	Ratio: 3.083	Gaps: 2							
Percent Similarity:	60.000	Percent Identity:	41.429						
alignment_block:									
US-09-215-435-177 x HSAIMICR	..								
Align. seg 1/1 to: HSAIMICR	from: 1 to: 1221								
24 GlyGly.....	ValAsnIleIlelaGlulysIleCysGlyAspLeuLy 38								
694 GGTGGCAACTGGTAACTGAAGTCACCAAGAAA.....	GA 728								
38 sAspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55									
729 AGATTCTCGCAGCTGGCTACTCGCCGGTCCCTGCATGGGAATGACCA 778									
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71									
779 GCAGGTATTTCATAATGGTACATCCATCGGCTGTGAGACTTTCACAGTAC 828									
72 SerGlyCysAsnGlyAsnLeuAsnAsnPhelysLeuLysIleGluArgGl 88									
829 GCGGCTCATGGGCAACGGTAACAACCTTCGTCACAGAAAAGGAGTGCT 878									
88 uValAlaCys 91									
879 CGAGACCTGC 888									
seq_name: gb_pat2:105217									
seq_documentation_block:									
LOCUS	105217	1232 bp							
DEFINITION	Sequence 7 from Patent EP 0255011.								
ACCESSION	105217								
VERSION	105217.1	GI:591138							
KEYWORDS									
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	1 (bases 1 to 1232)								
AUTHORS	KaumeYer,J.F., Kotick,M.P. and Polazzi,J.O.								
TITLE	Human inter-alpha-trypsin inhibitor gene								
JOURNAL	Patent: EP 0255011-A2 7 03-FEB-1988;								
FEATURES	Location/Qualifiers								
source	1..1232								
	/organism="unknown"								
BASE COUNT	300 a	323 c	365 g	244 t					

Thu Nov 15 10:51:52 2001

ORIGIN

```

alignment_scores:      Quality: 129.50      Length: 70
                       Ratio: 3.083         Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x I05217 ..
Align seg 1/1 to: I05217 from: 1 to: 1232

24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
   |||||  |||:|||||:|||||:|||||:|||||:|||||:|||||:
699 GGTGGCAACTGGTAAGTCAAGTCACCAAGAAA.....GA 733

38 sAspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
   |||  |||:|||||  :||  |||  |||  |||  |||  |||  |||
734 AGATTCTGCCAGCTGGGCTACTCGGCCGGTCCCTGCATGGGAATGACCA 783

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||
784 GCAGGTATTCTATATATGGTACATCCATGCCCTGTGAGACTTCCAGTAC 833

72 SerGlyCysAsnGlyAsnLeuAsnAsnAsnLysLeuLysIleGluArgGl 88
   :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||
834 GCGGGCTGCATGGCAACGGTAACAACTTCGTCAACAGAAAAGGAGTGTCT 883

88 uValAlaCys 91
   :|||
884 GCAGACCTGC 893

```


OM of: US-09-215-435-177 to: N_Geneseq_0601.* out_format : pfs
Date: Nov 15, 2001 4:26 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=7.000 -YGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@Cgnl_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-215-435-177
Query length: 99
Database: N_Geneseq_0601.*
Database sequences: 730101
Database length: 313950809
Search time (sec): 1043.840000

score_list:	Sequence	Strd Orig	zScore	EScore	Len	Documentation
	/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97612			539.00	1246.72	1.8e-61
	/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97738			539.00	1223.04	3.7e-60
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ00455			427.00	991.25	3.0e-47
	/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ51569			437.00	991.10	3.1e-47
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79088			140.00	312.99	1.8e-09
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79088			138.00	308.30	3.3e-09
	/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ32013			133.00	250.18	5.8e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ30007			133.00	250.18	5.8e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79087			131.50	291.77	2.8e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ35164			131.00	296.98	1.4e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ40464			129.50	288.97	4.0e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ81432			129.50	278.40	1.5e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ78950			129.50	278.21	1.6e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ77939			129.50	276.77	1.9e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1990.DAT:AAQ06878			127.00	291.72	2.8e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46936			127.00	283.66	7.9e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79081			126.00	282.52	9.1e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ15879			125.00	285.49	6.2e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79079			124.00	286.19	5.7e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79080			124.00	277.83	1.7e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79083			124.00	276.22	2.0e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ90469			124.00	267.88	6.0e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46935			123.00	274.29	2.6e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79077			123.00	282.28	9.4e-08
	/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ86987			121.00	266.01	7.6e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46934			120.00	267.26	6.5e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ31001			120.00	258.82	1.9e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46933			119.00	264.92	8.7e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ29024			118.50	268.95	5.2e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ29017			118.00	271.24	3.9e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ35169			118.00	267.90	6.0e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35175			118.00	264.52	9.2e-07
	/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ72238			118.00	252.49	4.3e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ10945			117.00	252.27	4.4e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ39947			117.00	249.15	6.6e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1991.DAT:AAQ10992			116.00	256.59	2.5e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1991.DAT:AAQ10993			116.00	252.69	4.2e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ51316			116.00	249.86	6.0e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ51317			116.00	249.86	6.0e-06
	/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ81396			116.00	249.68	6.2e-06

/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46644 + 116.00 249.50 6.3e-06
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46931 + 116.00 249.50 6.3e-06
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ16056 + 116.00 248.27 7.4e-06
/cgnl_9/gcgdata/geneseq/geneseq/NA1989.DAT:AAQ90108 + 116.00 245.23 1.1e-05
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ32167 + 116.00 245.23 1.1e-05
seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97612

seq_documentation_block:
ID_AAX97612 standard; DNA; 526 BP.
XX
AC_AAX97612;
XX
DT_13-SEP-1999 (first entry)
XX
DE_Extended human secreted protein coding sequence, SEQ ID NO. 76.
XX
KW_Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW_cellular differentiation; immune system regulator; anti-inflammatory;
KW_haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW_reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW_genetic disease; ss.
XX
OS_Homo sapiens.
XX
PN_WO9931236-A2.
XX
PD_24-JUN-1999.
XX
PF_17-DEC-1998; 98WO-IB02122.
XX
PR_10-AUG-1998; 98US-00961116.
PR_17-DEC-1997; 97US-0069957.
PR_09-FEB-1998; 98US-0074121.
PR_13-APR-1998; 98US-0081563.
XX
(GEST) GENSET.
XX
Bougueleret L, Duclert A, Dumas Milne Edwards J;
WPI; 1999-385906/32.
DR_P-PSDB; AAY35928.
XX
New isolated human secreted proteins
XX
Claim 1; Page 210; 516pp; English.
XX
This sequence represents an extended human secreted protein coding
sequence of the invention. The secreted proteins can be used in treating
or controlling a variety of human conditions. The secreted proteins may
act as cytokines or may affect cellular proliferation or differentiation
or may act as immune system regulators, haematopoiesis regulators, tissue
growth regulators, regulators of reproductive hormones or cell movement
or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
tumour inhibition activity. The DNAs can be used in forensic procedures
to identify individuals or in diagnostic procedures to identify
individuals having genetic diseases resulting from abnormal expression of
the genes corresponding to the extended cDNAs. They are also useful for
constructing a high resolution map of the human chromosomes. They can
also be used for gene therapy to control or treat genetic diseases.
SQ Sequence 526 BP; 162 A; 100 C; 99 G; 165 T; 0 other;

alignment_scores:
Quality: 539.00 Length: 99
Ratio: 5.444 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-177 x AAX97612 ..
Align seg 1/1 to: AAX97612 from: 1 to: 526

us-09-215-435-177.rng

Thu Nov 15 10:51:53 2001

CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX
SQ Sequence 539 BP; 165 A; 103 C; 97 G; 162 T; 12 other;

alignment_scores: Quality: 529.00 Length: 99
 Ratio: 5.343 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.980

alignment_block:

US-09-215-435-177 x AAX97738 ..

Align seg 1/1 to: AAX97738 from: 1 to: 539

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
|||||
32 ATGAAGTCTGCCAAGCTGGGATTTCTTAAGATTTCTCATCTTCTGCTC 81
|||||
17 rLeuAsnThrLeuLeuGlyGlyValAsnLysLysIleAlaGluLysIleC 34
|||||
82 ATTGAATACCCCTGTTATTGGGTGCTGTTAATAAAATTGCGAGAAGATAT 131
|||||
34 ysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
|||||
132 GTGGAGACCTCAAGATCCCTGCAAAATTTGGACATGAATTTTGGAACTGC 181
|||||
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
|||||
182 TATGAAGTTACCTTATAGATATTTCTACACAGAACCTCCAAAGATGTGA 231
|||||
67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeu 84
|||||
232 AACCTTTTGTCTCTCCAGCTGTAATGGCAACCTTACAACTTCNAGCTTA 281
|||||
84 ysIleGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
|||||
282 AATAGACACCTGAAGTAKCCTGCTGTGCAAAATACAAACCCACCGAGG 328
|||||

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT: AAC00455

seq_documentation_block:

ID AAC00455 standard; cDNA; 267 BP.

XX AAC00455;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 453.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
|||||
22 ATGAAGTCTGCCAAGCTGGGATTTCTTAAGATTTCTCATCTTCTGCTC 71
|||||
17 rLeuAsnThrLeuLeuGlyGlyValAsnLysLysIleAlaGluLysIleC 34
|||||
72 ATTGAATACCCCTGTTATTGGGTGCTGTTAATAAAATTGCGAGAAGATAT 121
|||||
34 ysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
|||||
122 GTGGAGACCTCAAGATCCCTGCAAAATTTGGACATGAATTTTGGAACTGC 171
|||||
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
|||||
172 TATGAAGTTACCTTATAGATATTTCTACACAGAACCTCCAAAGATGTGA 221
|||||
67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeu 84
|||||
222 AACCTTTTGTCTCTCCGGCTGTAATGGCAACCTTACAACTTCNAGCTTA 271
|||||
84 ysIleGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
|||||
272 AATAGACACCTGAAGTAGCTGTTGCAAAATACAAACCCACCGAGG 318
|||||

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT: AAX97738

seq_documentation_block:

ID AAX97738 standard; DNA; 539 BP.

XX AAX97738;

XX 13-SEP-1999 (first entry)

XX Extended human secreted protein coding sequence, SEQ ID NO. 303.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.

XX Homo sapiens.

XX WO9931236-A2.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-IB02122.

XX 10-AUG-1998; 98US-0096116.

XX 17-DEC-1997; 97US-0069957.

XX 09-FEB-1998; 98US-0074121.

XX 13-APR-1998; 98US-0081563.

XX (GEST) GENSET.

XX Bouqueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI; 1999-385906/32.

XX P-PSDB; AAY36054.

XX New isolated human secreted proteins

XX Claim 1; Page 376; 516pp; English.

XX This sequence represents an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or

DR P-PSDB; AAG00449.
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 XX Claim 1; SEQ ID 453; 71pp + CD-ROM; English.
 XX
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 SQ Sequence 267 BP; 74 A; 57 C; 52 G; 84 T; 0 other;

alignment_scores:
 Quality: 427.00 Length: 79
 Ratio: 5.405 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.734
 alignment_block:
 US-09-215-435-177 x AAC00455 ...
 Align seg 1/1 to: AAC00455 from: 1 to: 267
 1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPheIlePheCys 17
 31 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTTCTTCATCTTCTGCTC 80
 17 rleuAsnThrLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
 81 ATTGAATACCCCTGTTATTGGTGGTGTAAATAAAATTCGGAGAAGATAT 130
 34 ySGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
 131 GTGGAGACTCAAAATCCCTGCAATTTGGACATCAATTTGGAACTGC 180
 51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
 181 TATGAAGTTCACCTTTAGATATTCTTACAACAGAACCTCCAAAAGATGTA 230

67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsn 79
 231 AACTTTTGTCTCTCCAGCTGTAATGGCAACCTTAAC 267
 seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AAx51569
 seq_documentation_block:
 ID AAX51569 standard; cDNA; 271 BP.
 XX
 XX AAX51569;
 AC
 XX
 XX 21-JUN-1999 (first entry)
 DT
 XX
 XX Human secreted protein 5' EST SEQ ID NO:148.
 DE
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX Homo sapiens.
 OS
 XX
 PN W09906549-A2.
 XX
 PD 11-FEB-1999.
 XX
 XX 31-JUL-1998; 98WO-1801231.
 PF
 XX
 XX 01-AUG-1997; 97US-0905279.
 PR
 XX
 XX (GEST) GENSET.
 PA
 XX
 XX Duclert A, Dumas Milne Edwards J, Lacroix B;
 PI WPI; 1999-153779/13.
 XX
 DR P-PSDB; AAY12791.
 DR
 XX
 XX New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from testis, ovary, uterus and spleen tissue
 PT
 XX
 PS Claim 1; Page 253-254; 522pp; English.
 XX
 CC AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12681 to
 CC AAY12913, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 SQ Sequence 271 BP; 75 A; 57 C; 55 G; 84 T; 0 other;

alignment_scores:
 Quality: 427.00 Length: 79
 Ratio: 5.405 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.734
 alignment_block:
 US-09-215-435-177 x AAX51569 ...
 Align seg 1/1 to: AAX51569 from: 1 to: 271
 1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPheIlePheCys 17
 32 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTTCTTCATCTTCTGCTC 81
 17 rleuAsnThrLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
 82 ATTGAATACCCCTGTTATTGGTGGTGTAAATAAAATTCGGAGAAGATAT 131
 34 ySGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
 132 GTGGAGACTCAAAATCCCTGCAATTTGGACATCAATTTTGAAGCTGC 181
 51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
 182 TATGAAGTTCACCTTTAGATATTCTTACAACAGAACCTCCAAAAGATGTA 231
 67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsn 79
 232 AACTTTTGTCTCTCCAGCTGTAATGGCAACCTTAAC 268

Thu Nov 15 10:51:53 2001

us-09-215-435-177.rng

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT: AAT79088
seq_documentation_block:
ID AAT79088 standard; DNA: 467 BP.
XX AC AAT79088;
XX DT 12-NOV-1997 (first entry)
XX DE Insert from plasmid pHH341.
XX KW Trypsin inhibitor; Kunitz domain; protease: active site; elastase;
XX KW neutrophil; disease; modification; site directed mutagenesis; ss.
XX OS Synthetic.
XX FH Key
XX CDS 7..441
XX FT /*tag= a
XX FT /product= elastase inhibitor Ep7-d21-RPDF-52-55
XX FT sig_peptide 7..66
XX FT /*tag= b
XX FT mat_peptide 67..438
XX FT /*tag= c
XX PN JP09124700-A.
XX PD 13-MAY-1997.
XX PF 07-NOV-1995; 95JP-0288527.
XX PR 07-NOV-1995; 95JP-0288527.
XX PA (GREC ) GREEN CROSS CORP.
XX DR WPI: 1997-316576/29.
XX CC P-PSDB; AAW25938.
XX PT New protease inhibitor - useful for treating diseases involving
XX CC elastase
XX PS Disclosure; Fig 38; 37pp; Japanese.
XX CC This is the nucleotide sequence of the insert for the plasmid pHH341
XX CC which contains the sequence encoding the novel elastase specific
XX CC inhibitor Ep7-d21-RPDF-52-55 (AAT79088). The inhibitor coding sequence
XX CC is linked downstream of the yeast invertase (SUC2) signal peptide
XX CC sequence. The modified protease inhibitors are targeted to the protease
XX CC elastase, especially from neutrophils and can be used to treat diseases
XX CC associated with elastase. Modifications of the active site were done by
XX CC site directed mutagenesis.
XX SQ Sequence 467 BP; 90 A; 121 C; 132 G; 124 T; 0 other;

alignment_scores:
Quality: 140.00 Length: 83
Ratio: 2.979 Gaps: 2
Percent Similarity: 56.627 Percent Identity: 40.964

alignment_block:
US-09-215-435-177 x AAT79088 ..

Align seg 1/1 to: AAT79088 from: 1 to: 467
9 LeuLeuArgPhePhePhePheCysSerLeuAsnThrLeuLeuLeuGlyL 25
|||||: |||: |||
13 CTTTTCGAAGCTTCCTTTTC.....CTTTTCGCTGGTTT 47
25 yValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysL 42
|||||: |||: |||
48 CGCTGCCAAGATATCTGCTAGA.....CCAGATTTCCTGTC 82

```

```

42 ysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPhe 58
::||| :: ||| ||| ::||| ||| ||| ||| ||| |||
83 AACTGGCTACTCGCGCGTCCCTGCTGCTATGTTCTCTAGGATATTC 132
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
59 TyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 75
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 TATAATGGTACATCCATGCCCTGTCAGACTTTCGTGTACGGCGGATGCAT 182
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
75 nGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 GGGCAACGGTAACTTCGTCACAGAAAGAGGTGCTGTGCAGACCTGC 231
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT: AAT79086
seq_documentation_block:
ID AAT79086 standard; DNA: 467 BP.
XX AC AAT79086;
XX DT 12-NOV-1997 (first entry)
XX DE Insert from plasmid pHH337.
XX KW Trypsin inhibitor; Kunitz domain; protease: active site; elastase;
XX KW neutrophil; disease; modification; site directed mutagenesis; ss.
XX OS Synthetic.
XX FH Key
XX CDS 7..441
XX FT /*tag= a
XX FT /product= elastase inhibitor
XX FT sig_peptide 7..66
XX FT /*tag= b
XX FT mat_peptide 67..438
XX FT /*tag= c
XX PN JP09124700-A.
XX PD 13-MAY-1997.
XX PF 07-NOV-1995; 95JP-0288527.
XX PR 07-NOV-1995; 95JP-0288527.
XX PA (GREC ) GREEN CROSS CORP.
XX DR WPI: 1997-316576/29.
XX CC P-PSDB; AAW25936.
XX PT New protease inhibitor - useful for treating diseases involving
XX CC elastase
XX PS Disclosure; Fig 30; 37pp; Japanese.
XX CC This is the nucleotide sequence of the insert for the plasmid pHH337
XX CC which contains the sequence encoding the novel elastase specific
XX CC inhibitor Epi-d21-RPDF-52-55 (AAT79080). The inhibitor coding sequence
XX CC is linked downstream of the yeast invertase (SUC2) signal peptide
XX CC sequence. The modified protease inhibitors are targeted to the protease
XX CC elastase, especially from neutrophils and can be used to treat diseases
XX CC associated with elastase. Modifications of the active site were done by
XX CC site directed mutagenesis.
XX SQ Sequence 467 BP; 90 A; 123 C; 130 G; 124 T; 0 other;

alignment_scores:
Quality: 138.00 Length: 83
Ratio: 3.000 Gaps: 2
Percent Similarity: 55.422 Percent Identity: 40.964

```


KW		ss.
XX	Synthetic.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	1..441	
FT	/tag= a	
FT	/note= "no stop codon given"	
XX		
PN	WO9829453-A1.	
PN		
PD	09-JUL-1998.	
XX		
PF	05-JAN-1998; 98WO-JP00002.	
PR	27-DEC-1996; 96JP-0359053.	
PA	(MOCH) MOCHIDA PHARM CO LTD.	
PI	Hasegawa T, Kuriyama S;	
DR	WPI; 1998-388051/33.	
DR	P-PSDB; AAW69522.	
PT		
PT	Drugs containing peptide(s) with specific affinity to phospholipid(s) - such as phosphatidyl serine, for treatment of blood coagulation, inflammatory and immunological disorders	
PS	Example 10; Page 84; 117pp; Japanese.	
CC	The present invention describes drug compositions which contain as an active component a peptide which has specific affinity to particular phospholipids (such as phosphatidyl ethanolamine or phosphatidyl serine), especially to phospholipids which constitute a lipid bilayer of cellular cortex and of which the concentration in the bilayer increases in cells which are abnormal (e.g. through injury, denaturation or activation). In particular, the peptide contains a sequence having phospholipid affinity and a structure of formula (I); (A1)a-(A2)b-(A3)c,	
CC	where (A1) is one of two specific sequences (see AAW69516 and AAW69519), (A2) and (A3) are TRYLRHPQSWVHQIALR, LRYLRIHQSWVHQIALR (see AAW69517)	
CC	or MEVLGCEAQNLV (see AAW69518); a = 0-5; b = 1-5, and c = 0-5. Preferred are the formulae Al-A2-A3, A2-A2-A3, A2-A2-A2-A3 or A2-A2	
CC	(especially A2-A2-A3, A2-A2-A3 or A2-A2).	
CC	The sequence is linked to a peptide such as a blood factor, especially thrombo-modulin, urina-statin or membrane cofactor protein. The drugs are used for the treatment and prevention of diseases involving blood coagulation, inflammatory and immunological disorders. The present sequence encodes a protein from the present invention.	
XX	Sequence 441 BP; 99 A; 119 C; 133 G; 90 T; 0 other:	
SQ		
alignment_scores:		
Quality:	129.50	Length: 70
Ratio:	3.083	Gaps: 2
Percent Similarity:	60.000	Percent Identity: 41.429
alignment_block:		
US-09-215-435-I77 x AAIV40046 ..		
Align seg 1/1 to: AAIV40046 from: 1 to: 441		
24 GlyLeu.....ValAsnLysIleAlaGlutylSerCysGlyAspLeuLy 38 : :::~::~:~::: 34 GTGGCACTACTGTAAAGTACCAAGA...GA 68 38 saspProCysLysLeuAspMetAsnPheCysTyrgluValHisP 55 : :::: 69 AGATTCTGCCCGTCGCATCTACTGCCCGTGCCCTGATGGCATGACCA 118 55 heArgTrvPheTrvArSnArgThrSrlYsaAgCysGluThrPheValphe 71		

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1988.DAT: AAN81432

seq_documentation_block:

ID AAN81432 standard; cdna; 1232 BP.

XX AC AAN81432;

DT DT 06-DEC-1990 (first entry)

DE DE Sequence encoding new fusion protein contg. alpha-l-microglobulin (AMG) and the HI-30 region of inter-alpha-trypsin inhibitor (III) light chain.

DE DE Serine protease; enzyme; pancreatitis; atherosclerosis;

KW KW Chronic inflammation; therapy; elastase; ss.

KW OS Homo sapiens.

XX Key Location/Qualifiers

CDS 51..1109

mat_peptide 108..656

product_AMG 666..1100

*tag= c

product-HI-30 921..926

*tag= d

Note-Differs from the protein sequence of HI-30 purified from urine

misc_feature 1077..1079

*tag= e

Note-Differs from the protein sequence of HI-30 purified from urine

EP255011-A.

03-FEB-1988.

20-JUL-1987; 87EP-0110461.

29-JUL-1986; 86US-0891469.

(MILE) MILES LABORATORIES INC.

Kaumeyer JF, Kotick MP, Polazzi JO;
WPI; 1988-030262/05.
P-PSDB; AAP81110.

New DNA sequence coding for fusion protein contg. alpha-microglobulin and inter-alpha-trypsin inhibitor, useful for treating excessive elastase prodn.

Disclosure ; p; English.

A cDNA sequence encoding AMG in sequence with the HI-30 coding region of the light chain is claimed. The cDNA also contains a 5'-sequence coding for a secretory leader. 2 Arg residues separate the HI-30 and AMG regions and preceding AMG is a signal sequence. ITI is a serine protease, potentially used for treating excessive release of hydrolytic enzymes, elastase, in conditions such as pancreatitis, atherosclerosis and chronic inflammation.


```

XX
SQ Sequence 1232 BP; 299 A; 325 C; 364 G; 244 T; 0 other;

alignment_scores:
  Quality: 129.50      Length: 70
  Ratio: 3.083         Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x AAN81432 ..

Align seg 1/1 to: AAN81432 from: 1 to: 1232

24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||:|||||:|||||
699 GGTGGCAACTGGTAACTGAGTCACCAAGAA.....GA 733

38 sAspProCysLysLeuAspMetAsnPhcGlySerCysTyrGluValHisP 55
:||||| |||:|||||: |||
734 AGATTCTCTGCCAGCTGGGCTACTCGCGCGTCCCTGCATGGGAATGACCA 783

55 heArgTyrPhetYrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||:|||||: |||
784 GCAGGTATTCTATAATGGTACATCCATGGCTGTGAGACTTTCCAGTAC 833

72 SerGlyCysAsnGlyAsnLeuAsnAsnPhcLysLeuLysIleGluArgG1 88
:||||| |||:|||||: |||
834 GCGGCTGCATGGGCAACGGTAACAACCTTCGTACAGAAAGGAGTGCT 883

88 uValAlaCys 91
:|||||
884 GCAGACCTGC 893

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT78950

seq_documentation_block:
ID AAT78950 standard; DNA; 1255 BP.
XX
AC AAT78950;
XX
DT 11-NOV-1997 (first entry)
DE
XX Anti-trypsin inhibitor UTI coding sequence.
XX
KW Trypsin inhibitor; Kunitz domain; protease; active site; elastase;
KW neutrophil; disease; modification; site directed mutagenesis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 681..1124
FT /tag= a
FT /product= protease inhibitor UTI

JP09124700-A.
XX
XX 13-MAY-1997.
XX
XX 07-NOV-1995; 95JP-0288527.
XX
XX 07-NOV-1995; 95JP-0288527.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX WPI; 1997-316576/29.
XX
XX P-PSDB; AAW25928.
XX
XX New protease inhibitor - useful for treating diseases involving
XX elastase
XX
XX Disclosure; Fig 4-8; 37pp; Japanese.
XX
```

```

CC This is the nucleotide sequence encoding the anti-trypsin inhibitor UTI.
CC which is a protein containing 2 Kunitz domains. The nucleotide sequence
CC was used to construct a novel protease inhibitor by replacing the active
CC site of the Kunitz domain 1 (see AAT79080, AAT79081 and AAT79083).
CC The modified protease inhibitors are targeted to the protease elastase,
CC especially from neutrophils and can be used to treat diseases associated
CC with elastase. Modifications of the active site were done by site
CC directed mutagenesis.
XX
SQ Sequence 1255 BP; 310 A; 329 C; 370 G; 246 T; 0 other;

alignment_scores:
  Quality: 129.50      Length: 70
  Ratio: 3.083         Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x AAT78950 ..

Align seg 1/1 to: AAT78950 from: 1 to: 1255

24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||:|||||:|||||
714 GGTGGCAACTGGTAACTGAGTCACCAAGAA.....GA 748

38 sAspProCysLysLeuAspMetAsnPhcGlySerCysTyrGluValHisP 55
:||||| |||:|||||: |||
749 AGATTCTCTGCCAGCTGGGCTACTCGCGCGTCCCTGCATGGGAATGACCA 798

55 heArgTyrPhetYrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||:|||||: |||
799 GCAGGTATTCTATAATGGTACATCCATGGCTGTGAGACTTTCCAGTAC 848

72 SerGlyCysAsnGlyAsnLeuAsnAsnPhcLysLeuLysIleGluArgG1 88
:||||| |||:|||||: |||
849 GCGGCTGCATGGGCAACGGTAACAACCTTCGTACAGAAAGGAGTGCT 898

88 uValAlaCys 91
:|||||
899 GCAGACCTGC 908

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC77939

seq_documentation_block:
ID AAC77939 standard; cDNA; 1444 BP.
XX
AC AAC77939;
XX
DT 08-FEB-2001 (first entry)
DE
XX Human cancer associated gene sequence SEQ ID NO:333.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
XX WO20005350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05882.
XX
XX 12-MAR-1999; 99US-0124270.
XX
```

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-587533/55.
P-PSDB; AAB43730.

Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer -

Claim 1; Page 885; 2352pp; English.

AAC77607 to AAC78448 encode the human cancer associated proteins given in AAB43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; tumorigenic; immunomodulatory; antiangiogenic; antiproliferative; antirheumatic; antidiabetic; antitubercular; antihypertensive; antibacterial; antiviral; antiinflammatory; antithyroidal; antiallergic; cardioprotective; coagulant; dermatological; neuroprotective; antipsoriatic and antiangiotensin. The neotropic; vasotrophic; antipsoriatic and antiangiogenic. The polynucleotide and polypeptides may be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions from polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostasis or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological diseases and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.

Sequence 1444 BP; 362 A; 383 C; 422 G; 277 T; 0 other;

alignment_scores:
Quality: 129.50 Length: 70
Ratio: 3.083 Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x AAC77939 ..

Align seg 1/1 to: AAC77939 from: 1 to: 1444

24 GlyGly.....ValAsnLysIleAlaGlutylCysGlyAspLeu 38
||||| |||||GA 917
883 GTGGCACTGTAACTAAGTACCAGAANA.....

38 saspProcysLysLeuAspMetAsnPheGlySerCystyrGluHis 55
|||||
918 AGATTCTTCGCATCGGCTACTCCGCCGTTCCGTCATGGAATGACCA 967

55 heArgTyrrPhetYrAsnArghThrSerLysArgCysGluThrpheValphe 71
|||||
968 GCAGGTATTCTATAATGTATCATCATCGGCTTGAGACATTTCCAGTAC 1017

72 SerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGI 88
|||||
1018 GGCGGCTGTCATGGCACGGTAACAACITTCGTACAGAAAAGGAGTGTCT 1067

88 uValAlaCys 91
|||||
1068 GCAGACCTGC 1077

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1990.DAT.AAQ06878
seg_documentation_block:

```
68 TTCTCGTTACTTCTACAACGGTACCTCTATGGCTTGGGAAACTTTCCAGT 117
71 heSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArg 87
   :::::|||||  |||||  |||||||||  ::  |||
118 ACGGTGGTTGCATGGGTACGGTAACAACTTCGTTACTGAAAAAGAAATGC 167
   :::::|||||  |||||  |||||||||  ::  |||
88 GluValAlaCys 91
   :::::|||||
168 CTGCAGACTTGC 179
```

THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-177 to: Issued_Patents_NA.* out_format : pfs
Date: Nov 15, 2001 3:58 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=rni -GAPOP=12_000
-GAPEXT=4_000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4_500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6_000 -FCGAPEXT=7_000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6_000 -DELEXT=7_000 -START=1 -MATRIX=biosum62
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177
Query length: 99
Database: Issued_Patents_NA.*
Database sequences: 351203
Database length: 11323899
Search time (sec): 228.030000

score_list:	Strd Orig	zScore	EScore	Len	Documentation
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-6 +	131.00	293.43	1.1e-08	1455	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-6588-17 +	127.00	296.56	7.2e-09	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-147-710-1 +	124.00	280.31	5.8e-08	979	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-458-090-1 +	124.00	280.31	5.8e-08	979	1
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-457-887-1 +	124.00	280.31	5.8e-08	979	1
/cgnl_7/ptodata/1/ina/6A.COMB.seq:US-08-817-145-2 +	124.00	278.73	7.1e-08	1140	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-6588-15 +	123.00	286.71	2.6e-08	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-6588-13 +	120.00	279.32	6.6e-08	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-6588-11 +	119.00	276.86	9.0e-08	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-135 +	118.00	279.78	6.2e-08	249	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-2 +	116.00	268.16	2.8e-07	474	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-4 +	116.00	264.20	4.6e-07	693	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-026-145-1 +	116.00	261.17	6.7e-07	928	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-446-646-8 +	116.00	261.17	6.7e-07	928	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-6588-7 +	116.00	280.98	6.9e-07	945	1
/cgnl_7/ptodata/1/ina/6A.COMB.seq:US-08-854-764-1 +	116.00	259.73	8.1e-07	1065	1
/cgnl_7/ptodata/1/ina/6B.COMB.seq:PCR-US95-09377-1 +	116.00	259.73	8.1e-07	1065	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-09-054-782-1 +	116.00	256.66	1.2e-06	1431	1
/cgnl_7/ptodata/1/ina/backfile1.seq:5466783-1 +	116.00	256.66	1.2e-06	1431	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-6 +	115.00	263.00	5.3e-07	614	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-147-710-14 +	114.00	274.22	1.3e-07	165	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-458-090-14 +	114.00	274.22	1.3e-07	165	1
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-457-887-14 +	114.00	274.22	1.3e-07	165	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-70 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-74 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-70 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-74 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-70 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-74 +	113.00	262.23	5.9e-07	412	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-58 +	111.00	257.15	1.1e-06	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-58 +	111.00	257.15	1.1e-06	418	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-58 +	111.00	257.15	1.1e-06	418	1
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-829-876-90 +	110.00	254.04	1.7e-06	445	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-685-660A-4 +	110.00	248.48	3.4e-06	759	1
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-974-196-4 +	110.00	248.48	3.4e-06	759	1
/cgnl_7/ptodata/1/ina/6B.COMB.seq:US-09-013-896A-1 +	110.00	240.65	9.4e-06	1610	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-985-692-1 +	107.50	249.02	3.2e-06	399	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-155-331-1 +	107.50	249.02	3.2e-06	399	1
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-424-022-1 +	107.50	249.02	3.2e-06	399	1
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-424-017B-1 +	107.50	249.02	3.2e-06	399	1

/cgnl_7/ptodata/1/ina/PTCUS.COMB.seq:PCT-US93-11696-1 + 107.50 249.02 3.2e-06 399
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-155-331-12 + 107.50 225.76 6.3e-05 3725
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-424-022-12 + 107.50 225.76 6.3e-05 3725
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-424-017B-12 + 107.50 225.76 6.3e-05 3725
/cgnl_7/ptodata/1/ina/PTCUS.COMB.seq:PCT-US93-11696-12 + 107.50 225.76 6.3e-05 3
seq_name: /cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-6

seq_documentation_block:
: Sequence 6, Application US/08358160
: Patent No. 5663143
: GENERAL INFORMATION:
: APPLICANT: LEY, Arthur C.
: APPLICANT: LADNER, Robert C.
: APPLICANT: GUTERMAN, Sonia K.
: APPLICANT: ROBERTS, Bruce L.
: APPLICANT: MARKLAND, William
: APPLICANT: KENT, Rachel B.
: TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
: TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
: NUMBER OF SEQUENCES: 234
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W. Suite 300
: CITY: Washington
: STATE: District of Columbia
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/358,160
: FILING DATE: 16-DEC-1994
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/133,031
: FILING DATE: 13-OCT-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/009,319
: FILING DATE: 26-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/664,989
: FILING DATE: 01-MAR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/487,063
: FILING DATE: 02-MAR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/240,160
: FILING DATE: 02-SEP-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Cooper, Iver P.
: REGISTRATION NUMBER: 28,005
: REFERENCE/DOCKET NUMBER: LEY=1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1455 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: synthetic DNA fragment
US-08-358-160-6

alignment_scores:
Quality: 131.00 Length: 85

us-09-215-435-177.rni

Thu Nov 15 10:51:53 2001

```

Ratio: 2.911                Gaps: 2
Percent Similarity: 52.941   Percent Identity: 37.647

alignment_block:
US-09-215-435-177 x US-08-358-160-6 ..
Align seg 1/1 to: US-08-358-160-6 from: 1 to: 1455

22 LeuLeuGlyGlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLys 38
||||| : : : : : |||
10 TTTATTTCGCAATTCCTTAGTTGTTCTTCTATTCTGCGCCAAGA 59
||||| : : : : : |||
38 sAspProCysLysLeuAspMetAsnPhelSerCysTyrGluValHisPhe 55
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
60 AGACTCTGCCAGCTGGGCTACTCGCGCGTCCCTGCATCGGATGACCA 109
||||| : : : : : |||
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| : : : : : |||
110 GCAGGTATTTCTATATGCTACATCCATGGCCTGTGAGACTTTCCAGTAC 159
||||| : : : : : |||
72 SerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGlu.... 86
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
160 GCGGCTGCTGCGCAACGTAACAACCTTCGTCACAGAAAGAGAGTGTCT 209
||||| : : : : : |||
87 .....ArgGluVal.....AlaCysValA 93
||||| : : : : : |||
210 GCAGACTGCGCACTGCTGGCGCGCTGAAACTGTGTAAAGTTGTTAG 259
||||| : : : : : |||
93 laLys 94
|||||
260 CAAA 264

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-321-658B-17

seq_documentation_block:
; Sequence 17, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5576294o No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693,210-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: synthetic
FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
US-08-321-658B-17

alignment_scores:
Quality: 127.00      Length: 63
Ratio: 3.175        Gaps: 0
Percent Similarity: 63.492   Percent Identity: 41.270

alignment_block:
US-09-215-435-177 x US-08-321-658B-17 ..
Align seg 1/1 to: US-08-321-658B-17 from: 1 to: 418

29 IleAlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMe 45
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
212 ATGGCTGAGAGATGGAGAGAGAGAACGCCAGATTCTGCTTTTGAAGA 261
||||| : : : : : |||
45 tAsnPhelGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArg 62
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
262 AGATCCTGGAATATGTAAGCTCTGTTATATCAGGTATTTTATACAAATC 311
||||| : : : : : |||
62 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 78
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
312 AGACAAACACAGTGTGAACGTTTCAAGTATGTTGATGCGAGGGAATATG 361
||||| : : : : : |||
79 AsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 91
||||| : : : : : |||
362 AACAAATTTAAGACACTGGAAGAATGCAAGAACATTTGT 400
||||| : : : : : |||

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-147-710-1

seq_documentation_block:
; Sequence 1, Application US/08147710
; Patent No. 5455338
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,710
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E

```


us-09-215-435-177.rni

Thu Nov 15 10:51:53 2001

```

seq_documentation_block:
; Sequence 1, Application US/08457887
; Patent No. 5914315
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kistler, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457.887
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-14D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 979 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Placenta
; IMMEDIATE SOURCE:
; CLONE: J-2-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 39..746
; US-08-457-887-1

alignment_scores:
Quality: 124.00 Length: 86
Ratio: 2.385 Gaps: 1
Percent Similarity: 60.465 Percent Identity: 31.395

alignment_block:
US-09-215-435-177 x US-08-457-887-1 ..
Align seg 1/1 to: US-08-457-887-1 from: 1 to: 979

18 LeuAsnThrLeuLeuLeu.....GlyGlyValAsnLysIleAl 30
||||| |||||||
63 CTGTCGATTCGCTGCTTTCTCGAGGAGGCTGCACCTGGCGGATGCTGC 112
||||| |||||||
30 aGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsn 47
||||| |||||||
113 TCAGGAGCCCAACAGAAATACCGGAGATCTGCTCGGCCCTAGACT 162
||||| |||||||
47 heGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSer 63
||||| |||||||
163 ACGGACCTCCGGGCCCTACTTCTCCGTTACTACTACGACAGGTACAG 212
||||| |||||||

seq_documentation_block:
; Sequence 2, Application US/08817145
; Patent No. 6025329
; GENERAL INFORMATION:
; APPLICANT: UTSUMI, Jun
; APPLICANT: SUDO, Tetsuo
; APPLICANT: TANAKA, Yasuhiko
; APPLICANT: MATSUI, Mizuo
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR OPTHALMIC
; TITLE OF INVENTION: DISEASES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP.
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,145
; FILING DATE: 02-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-230P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..753
; US-08-817-145-2

alignment_scores:
Quality: 124.00 Length: 86
Ratio: 2.385 Gaps: 1
Percent Similarity: 60.465 Percent Identity: 31.395

alignment_block:
US-09-215-435-177 x US-08-817-145-2 ..
Align seg 1/1 to: US-08-817-145-2 from: 1 to: 1140

18 LeuAsnThrLeuLeuLeu.....GlyGlyValAsnLysIleAl 30
||||| |||||||

```



```
73 CTGTCGATTCTGCTCTTTCTGTCAGGAGGCTGCACCTGGCGGATGCTGC 122
30 aGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsn 47
123 TCAGGAGCAACAGGAATAACCGGAGATCTGCTCTGCCCCAGACT 172
47 heGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSer 63
173 ACGGACCTGCGCGGCGCTACTCTCTCCGTACTACTACGACAGGTACAG 222
64 LysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAs 80
223 CAGAGCTGCCGCGACTCTCTGTACGGGGCTGCGAGGCAACGCCAACAA 272
80 nPheLysLeuLysIleGluArgGluValAlaCysValAlaLysTyrLysp 97
273 TTCTACACCTGGGAGGCTTGCAGCAGATGCTGCGAGGATAGAAAGAG 322
97 roProArg 99
323 TTCCCAAA 330
```

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq.us-08-321-658B-15

```
seq_documentation_block:
; Sequence 15, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5576294o No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693.210-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: synthetic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
; FEATURE:
; NAME/KEY: sig_peptide
```

```
; LOCATION: 77..235
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
; US-08-321-658B-15

alignment_scores:
Quality: 123.00 Length: 63
Ratio: 3.075 Gaps: 0
Percent Similarity: 63.492 Percent Identity: 39.683

alignment_block:
US-09-215-435-177 x US-08-321-658B-15 ..
Align seg 1/1 to: US-08-321-658B-15 from: 1 to: 418.

29 IleAlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMe 45
:::|||||::: ||| ||| |||
212 ATGGCTGAGAGATTGGAGAAGAGAAAGCCAGATTTCTGCTTTTGAAGA 261
45 tAsnPheGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgT 62
::: ||| ||| ::||| ||| ||| |||
262 AGATCTGGAATATGTAAAGCTGATTATCAGGTATTTTATAACAATC 311
62 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 78
::: ||| ::||| ||| ::||| ||| ||| |||
312 AGACAAACACAGTGTGAACGTTTCAAGTATGGTGGATGACAGGGGCAAT 361
79 AsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
||||| ||| |||
362 AACAAATTTTGAGACACTGGAAGAATGCAAGAACATTGT 400

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq.us-08-321-658B-15

seq_documentation_block:
; Sequence 13, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5576294o No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693.210-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
```

Thu Nov 15 10:51:53 2001

us-09-215-435-177.rni

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: synthetic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
; US-08-321-658B-13

alignment_scores:
  Quality: 120.00      Length: 63
  Ratio: 3.077        Gaps: 0
  Percent Similarity: 61.905  Percent Identity: 39.683

alignment_block:
US-09-215-435-177 x US-08-321-658B-13 ..
Align seg 1/1 to: US-08-321-658B-13 from: 1 to: 418

29 IleAlaGluLysLeuLysGlyAspLeuLysAspProCysLysLeuAspMe 45
   :::::::::::::::::::: ||| ||| |||:::
212 ATGGCTGAGAGATTGGAGAGAGAACCCAGATTTCTGCTTTTGGAGA 261
   :::::::::::::::::::: ||| ||| |||:::
45 tAsnPheGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgT 62
   ::| ||| ||| ::::::::::::::::::::
262 AGATCCCTGGAATATGTAAGCTCGTATTATTCAGGTATTTTATAACAATC 311
   ::| ||| ||| ::::::::::::::::::::
62 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 78
   :::::::::::::::::::: ||| ::::::::::::::::::::
312 AGACAAACACAGTGTGAAGCTTCAAGTATGTTGATGCTGCGCAATATG 361
   :::::::::::::::::::: ||| ::|
79 AsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
   :::::::::::::::::::: ||| ::|
362 AACAAATTTGAGACACTGGAAGATGCAAGAACATTTGT 400
   ::| ||| ||| ::|

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-321-658B-11

seq_documentation_block:
; Sequence 11, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5576294 No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.658B
; FILING DATE: 12-OCT-1994

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693.210-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: synthetic/human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
; US-08-321-658B-11

alignment_scores:
  Quality: 119.00      Length: 63
  Ratio: 3.132        Gaps: 0
  Percent Similarity: 60.317  Percent Identity: 39.683

alignment_block:
US-09-215-435-177 x US-08-321-658B-11 ..
Align seg 1/1 to: US-08-321-658B-11 from: 1 to: 418

29 IleAlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMe 45
   :::::::::::::::::::: ||| ||| |||:::
212 ATGGCTGAGAGATTGGAGAGAGAACCCAGATTTCTGCTTTTGGAGA 261
   :::::::::::::::::::: ||| ||| |||:::
45 tAsnPheGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgT 62
   ::| ||| ||| ::::::::::::::::::::
262 AGATCCCTGGAATATGTCAGGTATTTATTCAGGTATTTTATAACAATC 311
   ::| ||| ||| ::::::::::::::::::::
62 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 78
   :::::::::::::::::::: ||| ::::::::::::::::::::
312 AGACAAACACAGTGTGAAGCTTCAAGTATGTTGATGCTGCGCAATATG 361
   :::::::::::::::::::: ||| ::|
79 AsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
   :::::::::::::::::::: ||| ::|
362 AACAAATTTGAGACACTGGAAGATGCAAGAACATTTGT 400
   ::| ||| ||| ::|

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-358-160-135

seq_documentation_block:
; Sequence 135, Application US/08358160
; Patent No. 5863143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300

```


us-09-215-435-177.rni

Thu Nov 15 10:51:53 2001

```
72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluAArgGlu 88
|||||
503 CAGGTATTTTATAACCAATCAGCAAAACAGTGTGAAGCTTCAAGTATC 552
|||||
72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluAArgGlu 88
|||||
553 GTGGATGCTGGGCAATATGACAAATTTTGAGACACTCGAGGAATGCAAG 602
|||||
89 ValAlaCys 91
|||
603 AACATTGT 611

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-026-145-1
seq_documentation_block:
; Sequence 1, Application US/08026145
; Patent No. 5378614
; GENERAL INFORMATION:
; APPLICANT: Petersen, Jens G. Litske
; APPLICANT: No. 5378614dfang, Ole Juul
; TITLE OF INVENTION: Method for Making TFFI Analogues
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 53786140 No. 5378614disk of No. 5378614th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6200
; CITY: New York
; STATE: N. Y.
; COUNTRY: United States of America
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/026.145
; FILING DATE: 19930302
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/828,920
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DK/90/00212
; FILING DATE: 17-AUG-1990
; APPLICATION NUMBER: DK 4080/89
; FILING DATE: 18-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Agriis, Cheryl H.
; REGISTRATION NUMBER: 34086
; REFERENCE/DOCKET NUMBER: 3321.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Synthetic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8..919
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 8..91
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 8..91

72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluAArgGlu 88
|||||
334 GTGGATGCTGGGCAATATGACAAATTTTGAGACACTCGAGGAATGCAAG 383
|||||
89 ValAlaCys 91
|||
384 AACATTGT 392

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-828-920A-4
seq_documentation_block:
; Sequence 4, Application US/07828920A
; Patent No. 5312736
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Jesper
; APPLICANT: No. 5312736dfang, Ole Juul
; TITLE OF INVENTION: Anticoagulant Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 53127360 No. 5312736disk of No. 5312736th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6200
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/828,920A
; FILING DATE: 19920127
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 4080/89
; FILING DATE: 18-AUG-1989
; APPLICATION NUMBER: WO PCT/DK90/00212
; FILING DATE: 17AUG1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T.
; REGISTRATION NUMBER: 30335
; REFERENCE/DOCKET NUMBER: 3287.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 105..656
; US-07-828-920A-4

alignment_scores:
Quality: 116.00 Length: 53
Gaps: 0
Ratio: 3.515
Percent Similarity: 62.264 Percent Identity: 43.396

alignment_block:
US-09-215-435-177 x US-07-828-920A-4 ..
Align seg 1/1 to: US-07-828-920A-4 from: 1 to: 693
39 AspProCysLysLeuAspMetAsnPhelGlySerCysTyrGluValHisPh 55
|||||
453 GATTCTGCTTTTGGGAAGAGGATCCTGGAATATGTCGAGGTATATATAC 502
|||||
```

```
;
; LOCATION: 92..919
; US-08-026-145-1

alignment_scores:
  Quality: 116.00      Length: 53
  Ratio: 3.515        Gaps: 0
  Percent Similarity: 62.264  Percent Identity: 43.396

alignment_block:
US-09-215-435-177 x US-08-026-145-1

Align seg 1/1 to: US-08-026-145-1 from: 1 to: 928

39 AspProCysLysLeuAspMetAsnPhelGlySerCysTyrGluValHisPh 55
||| ||| |||: |||: ||| |||
374 GATTCTGCTTTTGGAGAGAGATCTGGAATATGTCGAGGTTATATTAC 423

55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 72
|-||||| ||| |||: |||: ||| ||| ||| |||
424 CAGGTATTTTATAACAATCAGACAAAACAGTGTGAAAGGTTCAAGTATG 473

72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGlu 88
: ||| ||| ||| |||: ||| ||| ||| ||| ||| |||
474 GTGGATGCTGGCAATATGAACAATTTTGAGACACTCGAGGAATGCAAG 523

89 ValAlaCys 91
|||
524 AACATTGT 532

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-446-646-8

seq_documentation_block:
; Sequence 8, Application US/08446646
; Patent No. 5726038
; GENERAL INFORMATION:
; APPLICANT: Christiansen, Lars
; APPLICANT: Petersen, Jens G.
; TITLE OF INVENTION: A DNA Construct Encoding the YAP3 Signal
; TITLE OF INVENTION: Peptide
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57260380 No. 5726038disk of No. 5726038th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,646
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3987.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
;
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8..919
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 8..91
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 92..919
; US-08-446-646-8

alignment_scores:
  Quality: 116.00      Length: 53
  Ratio: 3.515        Gaps: 0
  Percent Similarity: 62.264  Percent Identity: 43.396

alignment_block:
US-09-215-435-177 x US-08-446-646-8

Align seg 1/1 to: US-08-446-646-8 from: 1 to: 928

39 AspProCysLysLeuAspMetAsnPhelGlySerCysTyrGluValHisPh 55
||| ||| |||: |||: ||| |||
374 GATTCTGCTTTTGGAGAGAGATCTGGAATATGTCGAGGTTATATTAC 423

55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 72
|-||||| ||| |||: |||: ||| ||| ||| |||
424 CAGGTATTTTATAACAATCAGACAAAACAGTGTGAAAGGTTCAAGTATG 473

72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGlu 88
: ||| ||| ||| |||: ||| ||| ||| ||| ||| |||
474 GTGGATGCTGGCAATATGAACAATTTTGAGACACTCGAGGAATGCAAG 523

89 ValAlaCys 91
|||
524 AACATTGT 532

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-321-658B-7

seq_documentation_block:
; Sequence 7, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 55762940 No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
```

Thu Nov 15 10:51:53 2001

REFERENCE/DOCKET NUMBER: 3693.210-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 945 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 365..538

US-08-321-658B-7

alignment_scores:

Quality: 116.00 Length: 53

Ratio: 3.515 Gaps: 0

Percent Similarity: 62.264 Percent Identity: 43.396

alignment_block:

US-09-215-435-177 x US-08-321-658B-7 ..

Align seg 1/1 to: US-08-321-658B-7 from: 1 to: 945

39 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 55

||| ||| |||::: ||| |||

371 GATTTCGCTTTTGGGAAGAAGATCCTGGAATATGTCGAGGTATATTAC 420

55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 72

||||| ||| |||::: ||| ||| ||| ||| ||| ||| ||| ||| |||

421 CAGGTATTTTATAACAAATCAGACAAACAGTGTGAACGTTTCAAGTATG 470

72 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 88

::||| ||| |||::: ||| ||| ||| ||| ||| ||| ||| ||| |||

471 GTGGATGCTGGGCAATATGAACAATTTTGAGACACTGGGAAGAATGCAAG 520

89 ValAlaCys 91

|||

521 AACATTTGT 529

[illegible]

us-09-215-435-177.rnmpm

Thu Nov 15 10:51:53 2001

```

seq_name: /cgnl_7/ptodata/1/pna/us6019_COMB.seq:US-60-197-873-5799

seq_documentation_block:
; Sequence 5799, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 5799
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..318
; NAME/KEY: sig_peptide
; LOCATION: 22..93
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.5999990463257
; OTHER INFORMATION: seq FFIFCSLNTLLG/GV
US-60-197-873-5799

alignment_scores:
Quality: 539.00 Length: 99
Ratio: 5.444 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177 x US-60-197-873-5799 ..
Align seg 1/1 to: US-60-197-873-5799 from: 1 to: 499

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
22 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTCTTCATCTTCGCTC 71
17 rLeuAsnThrLeuLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
72 ATGAATACCCCTGTTATTGGGTGGTGTAAATAAAATTCGCGAGAAGATAT 121
34 ysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
122 GTGGAGACCTCAAGATCCCTGCAAAATTTGGACATGAATTTTGGAGCTGC 171
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG1 67
172 TATGAAGTTCACCTTTAGATATTCTTACACAGAACCTCCAAAAGATCTGA 221
67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuL 84
222 AACTTTTGTCTCTCCGGCTGTAAATGGCAACCTTAACAACATTCGAAGCTTA 271
84 ysIleGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
272 AAATAGAAGCTGAAGTAGCTGTGTGTGCAAAATACAAACCCGAGG 318

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-76

seq_documentation_block:
; Sequence 76, Application US/09215435A
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bougueleret Lydie

```

```

; TITLE OF INVENTION: Extended CDNAS
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-2-9
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-4-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-8-10
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: Patent.pm
; SEQ ID NO 76
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..318
; NAME/KEY: sig_peptide
; LOCATION: 22..93
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq FFIFCSLNTLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 497..502
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 516..526
; US-09-215-435-76

alignment_scores:
Quality: 539.00 Length: 99
Ratio: 5.444 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177 x US-09-215-435-76 ..
Align seg 1/1 to: US-09-215-435-76 from: 1 to: 526

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
22 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTCTTCATCTTCGCTC 71
17 rLeuAsnThrLeuLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
72 ATGAATACCCCTGTTATTGGGTGGTGTAAATAAAATTCGCGAGAAGATAT 121
34 ysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
122 GTGGAGACCTCAAGATCCCTGCAAAATTTGGACATGAATTTTGGAGCTGC 171
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG1 67
172 TATGAAGTTCACCTTTAGATATTCTTACACAGAACCTCCAAAAGATCTGA 221
67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuL 84
222 AACTTTTGTCTCTCCGGCTGTAAATGGCAACCTTAACAACATTCGAAGCTTA 271
84 ysIleGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
272 AAATAGAAGCTGAAGTAGCTGTGTGTGCAAAATACAAACCCGAGG 318

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-303

seq_documentation_block:

```



```
; Sequence 303, Application US/09215435A
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bougueleret Lydie
; TITLE OF INVENTION: Extended cDNAs
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-2-9
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-4-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-8-10
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: Patent.pm
; SEQ ID NO 303
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..328
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.59999990463257
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 508..513
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 528..539
; US-09-215-435-303

alignment_scores:
    Quality: 529.00      Length: 99
    Ratio: 5.343        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 97.980

alignment_block:
US-09-215-435-177 x US-09-215-435-303  ..
Align seg 1/1 to: US-09-215-435-303 from: 1 to: 539
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
|||||
32 ATGAAGTCGCCAAGCTGGGATTTCTTAAAGATTCCTCATCTCTGCTC 81
|||||
17 rLeuAsnThrLeuLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
|||||
82 ATTGAATACCCCTGTTATTGGGTGGTGTAAATAAATTCGGGAGAAGATAT 131
|||||
34 ysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
|||||
132 GTGGAGACCTCAAAGATCCCTGCAATTTGGACATGAATTTTGAAGCTGC 181
|||||
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysGl 67
|||||
182 TATGAAGTTCACCTTAGATATTTCACACAGAACCTCCAAAGAAGATCA 231
|||||
67 uThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuL 84
|||||
232 AACTTTTGCTTCTCCAGCTGTATGGCAACCTTAACAACTTCAAGCTTA 281
|||||
84 ysIleGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
|||||
```

```
282 AATAGAACGTGAAGTAKCTGTGTGTCATAAATACAAACCCAGG 328
seq_name: /cgnl_7/ptodata/1/pna/US6006_COMB.seq:US-60-069-957-136

seq_documentation_block:
; Sequence 136, Application US/60069957
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: EXTENDED cDNAs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 381
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/069,957
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.019PR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 136:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..328
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 508..513
; FEATURE:
; NAME/KEY: polyA_a
; LOCATION: 528..539
; IDENTIFICATION METHOD: blastn
; US-60-069-957-136

alignment_scores:
    Quality: 529.00      Length: 99
    Ratio: 5.343        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 97.980

alignment_block:
US-09-215-435-177 x US-60-069-957-136  ..
Align seg 1/1 to: US-60-069-957-136 from: 1 to: 539
```

us-09-215-435-177.rnmp

Thu Nov 15 10:51:53 2001

```
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
|||||
32 ATGAAGTCTGCCAGCTGGGATTTCTTCTAAGATTCTTCATCTTCTGCTC 81
|||||
17 rLeuAsnThrLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
|||||
82 ATTGAATACCTCTTATTGGTGGTGTAAATAAATTCGGGAGAGATAT 131
|||||
34 ySgLYAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
|||||
132 GTGGAGACCTCAAGATCCCTGCAATTTGGACATGAATTTTGGAGAGCTGC 181
|||||
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG1 67
|||||
182 TATGAAGTTCACTTTAGATATTTCTACACAGAACCTTCAAGCTTA 231
|||||
67 uThrPheValPheSerGlyCysAsnGlyValAsnLysLeuL 84
|||||
232 AACTTTTGTCTTCTCCAGCTGTAAATGGCAACCTTAACAACCTTCAAGCTTA 281
|||||
84 ySileGluArgGluValAlaCysValAlaLysTyrLysProProArg 99
|||||
282 AAATAGAAGCTGAAGTAKCTCTGTGTGCAAAATACAAACCCAGGAGG 328
|||||

seq_name: /cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-148

seq_documentation_block:
; Sequence 148, Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; TITLE OF INVENTION: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knodde, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,279
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelisen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
;
US-08-905-279-148

alignment_scores:
Quality: 427.00 Length: 79
Ratio: 5.405 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.734

alignment_block:
US-09-215-435-177 x US-08-905-279-148 ..
Align seg 1/1 to: US-08-905-279-148 from: 1 to: 271
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
|||||
32 ATGAAGTCTGCCAGCTGGGATTTCTTCTAAGATTCTTCATCTTCTGCTC 81
|||||
17 rLeuAsnThrLeuLeuGlyGlyValAsnLysIleAlaGluLysIleC 34
|||||
82 ATTGAATACCTCTTATTGGTGGTGTAAATAAATTCGGGAGAGATAT 131
|||||
34 ySgLYAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
|||||
132 GTGGAGACCTCAAGATCCCTGCAATTTGGACATGAATTTTGGAGAGCTGC 181
|||||
51 TyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG1 67
|||||
182 TATGAAGTTCACTTTAGATATTTCTACACAGAACCTTCAAAAGATGTGA 231
|||||
67 uThrPheValPheSerGlyCysAsnGlyAsnLysLeuAsn 79
|||||
232 AACTTTTGTCTTCTCCAGCTGTAAATGGCAACCTTAAC 268

seq_name: /cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-516-448-1671

seq_documentation_block:
; Sequence 1671, Application US/09516448
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/516,448
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(349)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-516-448-1671

alignment_scores:
Quality: 368.00 Length: 68
Ratio: 5.662 Gaps: 0
Percent Similarity: 95.588 Percent Identity: 95.588

alignment_block:
US-09-215-435-177 x US-09-516-448-1671 ..
Align seg 1/1 to: US-09-516-448-1671 from: 1 to: 349
32 LysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheC1 48
|||||
23 AAGATATGTGGAGACCTCAAGATCCCTGCAATTTGGACATGAATTTGG 72
|||||
48 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 65
|||||
73 AAGCTGCTATGAAGTTCACTTTAGATATTTCTACACAGAACCTCCANNA 122
|||||
```

```

65 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 81
|||||
123 GATGTGAACCTTTGNCCTTCCTCGGCTGTAATGGCAACCTTTAAACAATTC 172

82 LysLeuLysIleGluAurGluValAlaCysValAlaLysTyrLysProPr 98
|||||
173 AAGCTTAAATANAACGTAAGTAGCCTGTGTTGCAAAATACAAACCACC 222

```

```

98 oArg 99
|||||
223 GAGG 226

```

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-833-381-1671

seq_documentation_block:

; Sequence 1671, Application US/09833381

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; PRIOR FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 1671

; LENGTH: 349

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)...(349)

; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1671

alignment_scores:

Quality: 368.00 Length: 68

Ratio: 5.662 Gaps: 0

Percent Similarity: 95.588 Percent Identity: 95.588

alignment_block:

US-09-215-435-177 x US-09-833-381-1671 ..

Align seg 1/1 to: US-09-833-381-1671 from: 1 to: 349

```

32 LysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheG1 48
|||||
23 AAGATATGTGGAGACCTCAAGATCCCTGCAAAATGGACATGAATTTGG 72

```

```

48 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 65
|||||
73 AAGCTGCTATGAAGTTCACCTTAGATATTTCTACAACAGAACCTCCANAA 122

```

```

65 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 81
|||||
123 GATGTGAACCTTTGNCCTTCCTCGGCTGTAATGGCAACCTTTAAACAATTC 172

```

```

82 LysLeuLysIleGluAurGluValAlaCysValAlaLysTyrLysProPr 98
|||||
173 AAGCTTAAATANAACGTAAGTAGCCTGTGTTGCAAAATACAAACCACC 222

```

```

98 oArg 99
|||||
223 GAGG 226

```

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-843-620-654

seq_documentation_block:

; Sequence 654, Application US/09843620

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

```

; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/843,620
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/200,099
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 1135
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 654
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(349)
; OTHER INFORMATION: n = A,T,C or G
US-09-843-620-654

```

alignment_scores:

Quality: 368.00 Length: 68

Ratio: 5.662 Gaps: 0

Percent Similarity: 95.588 Percent Identity: 95.588

alignment_block:

US-09-215-435-177 x US-09-843-620-654 ..

Align seg 1/1 to: US-09-843-620-654 from: 1 to: 349

```

32 LysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheG1 48
|||||
23 AAGATATGTGGAGACCTCAAGATCCCTGCAAAATGGACATGAATTTGG 72

```

```

48 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 65
|||||
73 AAGCTGCTATGAAGTTCACCTTAGATATTTCTACAACAGAACCTCCANAA 122

```

```

65 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 81
|||||
123 GATGTGAACCTTTGNCCTTCCTCGGCTGTAATGGCAACCTTTAAACAATTC 172

```

```

82 LysLeuLysIleGluAurGluValAlaCysValAlaLysTyrLysProPr 98
|||||
173 AAGCTTAAATANAACGTAAGTAGCCTGTGTTGCAAAATACAAACCACC 222

```

```

98 oArg 99
|||||
223 GAGG 226

```

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-188-162-1870

seq_documentation_block:

; Sequence 1870, Application US/60188162

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

; FILE REFERENCE: CL000518

; CURRENT APPLICATION NUMBER: US/60/188,162

; CURRENT FILING DATE: 2000-05-09

; NUMBER OF SEQ ID NOS: 5094

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1870

; LENGTH: 585

; TYPE: DNA

; ORGANISM: HUMAN

US-60-188-162-1870

alignment_scores:

us-09-215-435-177.rnps

Thu Nov 15 10:51:53 2001

```

Quality: 335.00      Length: 60
Ratio: 5.583         Gaps: 0
Percent Similarity: 100.000      Percent Identity: 98.333

alignment_block:
US-09-215-435-177 x US-60-188-162-1870/rev ..
Align seg 1/1 to reverse of: US-60-188-162-1870 from: 1 to: 585

39 AspProCysLysLeuAspMetAsnPhelSerCysTyrGluValHisPh 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 GATCCCTGCAATATGGACATGAATTTTGGAGCTGCTATGAAGTTCACT 207
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
206 TAGATATTCTACACAGACAGCTCCAAAGATGTGAACATTTCCTCTCT 157
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGlu 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
156 CCAGCTGTAAATGGCAACCTTACAACCTTCAAGCTTAAATAGAAGTGA 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
89 ValAlaCysValAlaLysTyrLysProPro 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 GTAGCTGTGTTGCAAAATACAAACCCG 77

seq_name: /cgnl_7/ptodata/1/pna/US095C_comb.seq:US-09-540-208-69337

seq_documentation_block:
; Sequence 69337, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; EARLIER FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 69337
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: rat00103589
; NAME/KEY: unsure
; LOCATION: 22, 255
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-208-69337

alignment_scores:
Quality: 251.00      Length: 71
Ratio: 4.328         Gaps: 0
Percent Similarity: 81.690      Percent Identity: 59.155

alignment_block:
US-09-215-435-177 x US-09-540-208-69337 ..
Align seg 1/1 to: US-09-540-208-69337 from: 1 to: 282

26 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 GTGGAGAGGTTGCTGCAAGGCTGCAAGAAATTCACGATCCTTGTC 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
42 sLeuAspMetAsnPhelGlySerCysTyrGluValHisPheArgTyrPhe 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 TTGTGAATGGATCCTGGCAGCTGCTATGAAGCCATTACAGATTTTCT 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

59 yrasnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 ACAACCAAACTCCAAAAGTGTGAGATTTCGTGTTCAGTGGTTGCAAT 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
76 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCysVa 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 GGTAACCCGTAATAACTCACTTAAATAAGACTGTGACGCTAATGTGTA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
92 lAlaLysTyrLys 96
|||||:|||||
201 TGAAGATAACAAG 213

seq_name: /cgnl_7/ptodata/1/pna/US091_comb.seq:US-09-131-380-189

seq_documentation_block:
; Sequence 189, Application US/09131380
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Garrow, Bonnie L.
; APPLICANT: Klemm, Juli D.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT FEMUR
; FILE REFERENCE: PZ-0013 US
; CURRENT APPLICATION NUMBER: US/09/131,380
; CURRENT FILING DATE: 1998-08-07
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/074,278
; EARLIER FILING DATE: 1998-02-10
; NUMBER OF SEQ ID NOS: 2199
; SOFTWARE: PERL Program
; SEQ ID NO 189
; LENGTH: 208
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 109, 117, 156, 167
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE:
; OTHER INFORMATION: 700292496H1
US-09-131-380-189

```

```

alignment_scores:
Quality: 210.00      Length: 67
Ratio: 4.038         Gaps: 1
Percent Similarity: 77.612      Percent Identity: 55.224

alignment_block:
US-09-215-435-177 x US-09-131-380-189 ..
Align seg 1/1 to: US-09-131-380-189 from: 1 to: 208

26 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2 GTGGAGAGGTTGCTGCAAGGCTGCAAGAAATTCACGATCCTTGTC 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
42 sLeuAspMetAsnPhelGlySerCysTyrGluValHisPheArgTyrPhe 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 TTGTGAATGGATCCTGCGAGCTGCTATGAAGCCATTACAGATTTTCT 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59 yrasnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 ACAACCAAACTCCCAAGCTGTCAGATTTTCGTGTTTCAGTGGTTGCA 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
75 nGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 TGTNACCGTGAATAAATTCATGCTTAAATAAGACTGTGACGTAATGTGT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: /cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-499-23002

seq_documentation_block:
; Sequence 23002, Application US/09540499
; GENERAL INFORMATION:
; APPLICANT: Seillamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MUSCULOSKELETAL SYSTEM TISSUE
; FILE REFERENCE: PD-1032 CIP
; CURRENT APPLICATION NUMBER: US/09/540,499
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 29520
; SOFTWARE: PERL Program
; SEQ ID NO 23002
; LENGTH: 208
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: rat00198915
; NAME/KEY: unsure
; LOCATION: 109, 117, 156, 167
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-499-23002

alignment_scores:
Quality: 210.00 Length: 67
Ratio: 4.038 Gaps: 1
Percent Similarity: 77.612 Percent Identity: 55.224

alignment_block:
US-09-215-435-177 x US-09-540-499-23002 ..

Align seg 1/1 to: US-09-540-499-23002 from: 1 to: 208

```
26 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2 GTGGAGAAGGTTGCTGCAAGGCTCTGCAAGAATTCACAGATCCTTGTC 51

42 sleuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 TTTGAAAATGGATCTGTCGAGCTGCTATGAAAGCCATTACAGATTTTCT 101

59 yrAsnArgThrSer.LysArgCysGluThrPheValPheSerGlyCysAs 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 ACAACCANACCTCCANCCAGTGTGAGATTTTCGTGTTTCAGTGGTGC 151

75 nGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 TGGTNACCGTAATAATTCATGCTTAAATAAGACTGTGACGTAATGTGT 200
```

seq_name: /cgnl_7/ptodata/1/pna/US6005_COMB.seq:US-60-055-705-189

seq_documentation_block:
; Sequence 189, Application US/60055705
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Argentine, Rebecca E.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT
; TITLE OF INVENTION: FEMUR
; NUMBER OF SEQUENCES: 1093
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO

; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/055,705
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PZ-0013 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 189:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 700292496H1
US-60-055-705-189

alignment_scores:
Quality: 210.00 Length: 67
Ratio: 4.038 Gaps: 1
Percent Similarity: 77.612 Percent Identity: 55.224

alignment_block:
US-09-215-435-177 x US-60-055-705-189 ..

Align seg 1/1 to: US-60-055-705-189 from: 1 to: 208

```
26 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2 GTGGAGAAGGTTGCTGCAAGGCTCTGCAAGAATTCACAGATCCTTGTC 51

42 sleuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 TTTGAAAATGGATCTGTCGAGCTGCTATGAAAGCCATTACAGATTTTCT 101

59 yrAsnArgThrSer.LysArgCysGluThrPheValPheSerGlyCysAs 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 ACAACCANACCTCCANCCAGTGTGAGATTTTCGTGTTTCAGTGGTGC 151

75 nGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCys 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 TGGTNACCGTAATAATTCATGCTTAAATAAGACTGTGACGTAATGTGT 200
```

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-480-902-6802

seq_documentation_block:
; Sequence 6802, Application US/09480902
; GENERAL INFORMATION:
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; APPLICANT: Tao, Nengbing
; APPLICANT: Warren, Wesley C.
; TITLE OF INVENTION: Nucleic Acid and Other Molecules Associated with Lactation and
; TITLE OF INVENTION: Muscle and Fat Deposition
; FILE REFERENCE: 10298/1
; CURRENT APPLICATION NUMBER: US/09/480,902
; CURRENT FILING DATE: 2000-01-11
; EARLIER APPLICATION NUMBER: USSN 60/115,707
; EARLIER FILING DATE: 1999-01-12

```
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6802
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 29-LIB34-029-Q1-EL-H1
US-09-480-902-6802

alignment_scores:
    Quality: 134.00      Length: 43
    Ratio: 4.467         Gaps: 0
    Percent Similarity: 69.767   Percent Identity: 55.814

alignment_block:
US-09-215-435-177 x US-09-480-902-6802 ..

Align seg 1/1 to: US-09-480-902-6802 from: 1 to: 404

39 AspProCysLysLeuAspMetAsnPhcGlySerCysTyrGlutValHisph 55
||| |||:::||||| ::| ||| |||
239 GATTCTCGCAGCTGGACTACTCACAAAGGCCTTCGCCGGGCGCTTTCAA 288
||| |||:::||||| |||||
55 eacTyrPheTyrAsnArgHrSerLysArgCysGluThrPheValPheS 72
||||| |||:::||||| |||||
289 GAGGTATTTCTATAACGGTACATCCATGCCCTGTGAGACCTTCTTATATG 338
||||| |||:::||||| |||||

72 erGlyCysAsnGlyAsnLeuAsnAsphe 81
||||| |||:::||||| |||||
339 CGCGTGTCATGGGAATGGCAACAATTC 367
```


us-09-215-435-177.rnpn

Thu Nov 15 10:51:54 2001

US-09-215-435-177 x US-08-849-406A-78
Align seg 1/1 to: US-08-849-406A-78 from: 1 to: 249
22 LeuLeuGlyGlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
10 CTTCTCTCCGCTATCTCTGGTGGTACCTTTCTATTCCGGCGCCCAAGCC 59
38 sAspProCysLysLeuAspMetAsnPhelGlySerCysTyrGluValHisP 55
60 TGACTTCTGCTCTCCGAGGAGGATCCGGGATTTGCCGCGGTATATTA 109
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
110 CGCGTTATTCTTATTAATAACAGACTAAGCAATCTGAGCGGTTCAAGTAT 159
72 SerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgG1 88
160 GGTGCTGCTAGGTAATGAACAACCTCGAGACTCTAGAGAGACTGTAA 209
88 uValAlaCys 91
210 GAACATATCT 219
seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-84
seq_documentation_block:
; Sequence 84, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-08-849-406A-84

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
110 GCAGGTATTCTATAATGGTACATCCATGGCCCTGTGAGACTTTCCAGTAC 159
72 SerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGlu..... 86
160 GCGCGCTGCATGGCGCAACGGTAACAACCTTCGTACAGAAAAGGAGTGCT 209
87ArgGluVal.....AlaCysValA 93
210 GCAGACCTGCCGCAACTGTGGCGCGCTGAACCTTCAAAAGTGTGTAG 259
93 lalys 94
260 CAAA 264
seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-78
seq_documentation_block:
; Sequence 78, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-08-849-406A-78
alignment_scores:
Quality: 118.00 Length: 70
Ratio: 2.950 Gaps: 0
Percent Similarity: 57.143 Percent Identity: 37.143
alignment_block:
.....


```

alignment_scores:
  Quality: 105.00      Length: 57
  Ratio: 3.387         Gaps: 0
  Percent Similarity: 54.386      Percent Identity: 36.842

alignment_block:
  US-09-215-435-177 x US-08-849-406A-84  ..

Align seg 1/1 to: US-08-849-406A-84 from: 1 to: 201

35 GlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCysCyst 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1  GCGCCGACCGTGCATCTCTGCTTCCTCGAGGAGNRTVVGSGGNNITGCR 50

51 rGluValHisPheArGlyrPheTyrrAsnArGthrSerLysArGcysGluT 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51  TGSTNWTITMNSCGTDTSTTCTATATAAACACGCTAAGCAATGTSWGV 100

68 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLys 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 NATTCVHATATGGTGTGTCVHGSGSTAATVBGAACAACCTCGAGACTCTA 150

85 IleGluArGluValAlaCys 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 GAAGAGTGTGAAGAACATATGT 171

seq_name: /cgn1_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-84

seq_documentation_block:
; Sequence 80, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEV=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

us-09-215-435-177.rnnpn

Thu Nov 15 10:51:54 2001

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-849-406A-1

alignment_scores:
  Quality: 99.50      Length: 77
  Ratio: 2.427       Gaps: 2
  Percent Similarity: 53.247   Percent Identity: 35.065

alignment_block:
US-09-215-435-177 x US-08-849-406A-1 ..
Align seg 1/1 to: US-08-849-406A-1 from: 1 to: 276

22 LeuLeuGlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
10 TTATTATTCGAATTCCTTTAGTGTGTTCTTCTATTCTGGCGCCGTC 59
38 sAspProCysLysLeuAspMetAsnPhgGlySerCysTyrGluValHisP 55
60 GGATTCTGTCTCGAGCCACCATACAGTGGCCCTGCAAGGCGCATCA 109
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
110 TCCGCTATTCTACAATGCTAAAGCAGCGCTGTGCCAGACCTTTGTATAC 159
72 SerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgL 88
160 GGTGGTTCCTGCTGCTAGCTAGCTACAACTTAA.....TCGGCCGA 200
88 u.ValAlaCysValAlaLysTyrLyspro 97
201 AGATTGCTGCTGCTACCTCGGTGGCGCG 229

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-11942

seq_documentation_block:
; Sequence 11942, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11942
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(538)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-11942

alignment_scores:
  Quality: 99.50      Length: 59
  Ratio: 3.109       Gaps: 2
  Percent Similarity: 54.237   Percent Identity: 38.983

alignment_block:
US-09-215-435-177 x US-09-922-340-11942 ..
Align seg 1/1 to: US-09-922-340-11942 from: 1 to: 6777

26 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 42
982 GTGGAAGAGGTGGTTCGAGAGGTGCTCTGAACAAGCCGAGACG..... 1026
42 sLeuAspMetAsnPhgGlySerCysTyrGluValHisPheArgTyrPheT 59
1027 .....GGCCGTCGCGAGCAATGATCTCCCGCTGCTACT 1060
59 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 75
1061 TTGATGTGACTGAAGGAGAGTGTGCCCATTCCTTTACCGCGGATGTGCG 1110
76 GlyAsnLeuAsnPhelLysLeuLysIleGluArgGluValAlaCysVa 92
1111 GGCAACCGGACAACTTT.....GACACAGAGAGTACTGTCAT 1148
92 lAla 93
1149 GGCC 1152

```

```

Align seg 1/1 to: US-09-922-340-11942 from: 1 to: 538

34 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPhgGlySerCy 50
331 TCGGTGGACCTGCCAGACACA.....GGACTCTG 359
50 sTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
360 CAAGGAGAGCATCCCGCGCTGTTACTACACCCCTTTCAGTGAACACTGC 409
67 luThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeu 83
410 CCCGCTTACCTATGTTGTTGTTATGGCAACAAGAACTTT..... 453
84 LysIleGluArgGluValAlaCysVal 92
454 .....GAGGAAGAGCAGCAGTGCCTC 474

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-149

seq_documentation_block:
; Sequence 149, Application US/09971392
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 6777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Template ID: 235164.25
US-09-971-392-149

alignment_scores:
  Quality: 92.50      Length: 68
  Ratio: 2.312       Gaps: 2
  Percent Similarity: 58.824   Percent Identity: 26.471

alignment_block:
US-09-215-435-177 x US-09-971-392-149 ..
Align seg 1/1 to: US-09-971-392-149 from: 1 to: 6777

```

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-9347
seq_documentation_block:
; Sequence 9347, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9347
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(426)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-9347

```

```

alignment_scores:
  Quality: 82.50      Length: 49
  Ratio: 3.056       Gaps: 2
  Percent Similarity: 55.102  Percent Identity: 40.816

alignment_block:
US-09-215-435-177 x US-09-922-340-9347 ..

Align seg 1/1 to: US-09-922-340-9347 from: 1 to: 426

34 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCys 50
||| ||||| |||
234 TCGGTGACCTGCCAGACACA.....GGACTCTG 262

50 sTyRGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 67
||| ||||| ||||| ||||| ||||| |||||
263 CAAGAGAGACATCCCGCGTGGTACTACAAACCCITCAGCGAACACTGCG 312

67 luThrPheValPheSerGlyCysAsnGlyAsn.LeuAsnAsnPhe 81
||| ||||| ||||| ||||| ||||| |||||
313 CCGGCTTACCTATGCTGTGTTGTACGNAACAANGAACAACTTT 357

```

```
seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-82
```

```

seq_documentation_block:
; Sequence 82, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-08-849-406A-82

```

```

alignment_scores:
  Quality: 81.00      Length: 51
  Ratio: 3.000       Gaps: 0
  Percent Similarity: 52.941  Percent Identity: 29.412

alignment_block:
US-09-215-435-177 x US-08-849-406A-82 ..

Align seg 1/1 to: US-08-849-406A-82 from: 1 to: 189

41 CysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTy 57
||| ||| ||||| ||||| ||||| |||||
22 TCGGCTTCAAAGCTNRTYVSGGTCNTGTRTGTGTTCTTCMNSCGTTD 71

57 rPheTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyC 74
||| ||||| ||||| ||||| ||||| |||||
72 STTCTTCAACATTTTCACGCGTCAGTGCWGVHATTCVHATACGGTGGTT 121

74 ysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAla 90
|| ||| ||||| ||||| ||||| |||||
122 GTVHGGSTAACSRGAACCGGTTTCAATCTCTAGAGGAATGTAAGAAGATG 171

91 Cys 91
|||
172 TGC 174

```

```

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-73

seq_documentation_block:
; Sequence 73, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA

```

us-09-215-435-177.rnnpn

Thu Nov 15 10:51:54 2001

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 11256
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-922-340-11256

alignment_scores:
    Quality: 76.00      Length: 34
    Ratio: 3.455        Gaps: 1
    Percent Similarity: 64.706    Percent Identity: 44.118

alignment_block:
    US-09-215-435-177 x US-09-922-340-11256 ..
    Align seg 1/1 to: US-09-922-340-11256 from: 1 to: 401

59 TyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 75
|||||
2 TACAACCCCTTCAGTGAACACTGCGCCGCTTTACCTATGTTGGTTGTTA 51
|||||
75 nGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCysV 92
|||||
52 TGGCAACAAGAACAACTTT.....GAGGAAGACACGACGAGTGCC 89
|||||
92 al 92
..
90 TC 91

seq_name: /cgnl_7/ptodata/1/pna/us08_NEW_COMB.seq:us-08-849-406A-71

seq_documentation_block:
; Sequence 71, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; seq_name: /cgnl_7/ptodata/1/pna/us09_NEW_COMB.seq:us-09-922-340-11256

seq_documentation_block:
; Sequence 11256, Application US/09922340

```

```
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8584 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 71:
US-08-849-406A-71

alignment_scores:
  Quality: 75.50      Length: 113
  Ratio: 1.425        Gaps: 4
  Percent Similarity: 46.903  Percent Identity: 28.319

alignment_block:
US-09-215-435-177 x US-08-849-406A-71 ..

Align seg 1/1 to: US-08-849-406A-71 from: 1 to: 8584

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhe..... 12
1030 CTGAAGACGAGACTGCTCAAAATCTGCTGAGGCTGTCATCGTTACTCT 1079
13 .....PheIlePheCysSerLeuAsnThrLeuL 22
1080 GACTTGGGAAGTGACTTCGACGTCGCTGTTTGCCATTCTCTAACTCTAC 1129
22 euLeu..Gly.....GlyValAsnLysIleAlaGluL 32
1130 TAACAACGGTTGTTGTTTCATCAACACATACCATCGCTTCTATCGCTGCTA 1179
32 ysIleCysGly.....AspLeuLysAspProCysLysLeuAspMet 45
1180 AGGAGGAGGCTTTCTTCCGACAGAGAGCTGCTGTGAACCTGCCAATC 1229
46 AsnPheGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgTh 62
1230 GTCAGAGGTCATGATGCTTCTTCCCAAGATGGCTTTCGACGCTGT 1279
62 rSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuA 79
1280 TAAGGTAAGTCGCTTCTTCCCATACGCTGTTGTCGAAGGTAACGGTA 1329
79 snAsnPhelLysLeuLysIleGluArgGluValAlaCys 91
1330 ACAAGTTCTACTCTGAGAGGAGGTGTAGAGAGTACTGT 1367

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-75

seq_documentation_block:
; Sequence 75, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
```

us-09-215-435-177.rnnpn

Thu Nov 15 10:51:54 2001

```

; PRIOR APPLICATION NUMBER: US 60/106588
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107089
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107034
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107035
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2194)
; US-09-830-446-15

alignment_scores:
  Quality: 68.00      Length: 58
  Ratio: 2.061       Gaps: 2
  Percent Similarity: 56.897   Percent Identity: 32.759

alignment_block:
  US-09-215-435-177 x US-09-830-446-15  ..

Align seg 1/1 to: US-09-830-446-15 from: 1 to: 2300

47 phegly.....SerCysTyrGluValHisPheArgTyrPheTyr... 59
|||||  ::::: |||||  ::::: |||||  ::::: |||||  ::::: |||||
1522 TTGGGAGGCTTATTAAAGTCCTTAGATCAAGATGAGATGATGTCATG 1571
60 .....AsnArgThrSerLysArgCysGluThrPheValPheS 72
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1572 TCTTGAATCCTAGATCATCGCTCAGGATACGCAACAGCTTCGTTTTT 1621
72 erGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGlu 88
::: |||||  |||  ||::: ||::: ||::: ||::: ||::: ||::: ||:::
1622 GTAGGCAACCGGTCCTTACCAACTATCAAAATATCAAAATATATCAAGAAC 1671
89 ValAlaCysValAlaLysTyrLys 96
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1672 AGGAACGTGAACGCAAAATATCGA 1695

```

```

seq_documentation_block:
; Sequence 8534, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FROM cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8534
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(418)
; OTHER INFORMATION: n = A,T,C or G
; US-09-922-340-8534

* alignment_scores:
  Quality: 75.00      Length: 33
  Ratio: 3.409       Gaps: 0
  Percent Similarity: 66.667   Percent Identity: 42.424

alignment_block:
  US-09-215-435-177 x US-09-922-340-8534  ..

Align seg 1/1 to: US-09-922-340-8534 from: 1 to: 418

59 TyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 75
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2 TACACCCCTTCAGTGAACACTGCGCGCTTTACCTATGGGGTGTGTTA 51
75 nGlyAsnLeuAsnPhelLysLeuLysIleGluArgGluValAlaCys 91
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
52 TGGCAACAAGAACACTTGTAGGAAAGACGACGTCCTCGAGTCCTGT 100

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-830-446-15

seq_documentation_block:
; Sequence 15, Application US/09830446
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.; OOMEN, Raymond P. and WANG, Joe
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there
; FILE REFERENCE: 77813-51
; CURRENT APPLICATION NUMBER: US/09/830,446
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/106034
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106039
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/10604
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106044
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106072
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106073
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106074
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106087
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106587
; PRIOR FILING DATE: 1998-11-02

```

OM of: US-09-215-435-177 to: EST:* out_format : pfs
Date: Nov 14, 2001 10:29 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framet-p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2118/app_query.fasta_1.1519
-DB=EST -OFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435_@CGL_1_6733 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-215-435-177
Query length: 99
Database: EST:*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 5896.450000

score_list:	Strd	Orig	zScore	EScore	Len	Documentation
gb_est45:AW303354	-	433.00	1006.35	6.2e-47	443	AW303354 xv18p01.x1 Soares.NFL
gb_est45:AW303354	-	324.00	746.24	1.9e-32	611	AK020315 Mus musculus adult mal
gb_est45:AW303354	-	235.00	537.19	8.4e-21	570	AZ763929 IM0559C1R Mouse 10kb
gb_est45:AW303354	-	132.00	295.94	2.3e-07	552	A1256949 LP04037.5prime LP Dros
gb_est45:AW303354	-	132.00	295.93	2.3e-07	496	A1015730 ot71b08.s1 Soares.tota
gb_est45:AW303354	-	129.50	293.43	3.2e-07	350	T74655 yc57ell.r1 Stratagene li
gb_est45:AW303354	-	129.50	293.43	3.2e-07	413	AA424642 zv91d02.r1 Soares.Nhm
gb_est45:AW303354	-	129.50	289.64	5.2e-07	511	AF063548 AF063548 Homo sapiens
gb_est45:AW303354	-	129.50	289.64	5.2e-07	511	A1106714 HA0104 Human fetal liv
gb_est45:AW303354	-	129.50	289.45	5.3e-07	521	A1207711 HA03093 Human fetal liv
gb_est45:AW303354	-	129.50	288.53	5.9e-07	571	A1207590 HA2954 Human fetal liv
gb_est45:AW303354	-	129.50	287.61	6.7e-07	626	AV658823 AV658823 GLC Homo sap
gb_est45:AW303354	-	129.50	287.32	6.9e-07	644	A1807643 wf49d04.x1 Soares.NFL
gb_est45:AW303354	-	129.50	287.05	7.2e-07	662	A1065076 HA0910 Human fetal liv
gb_est45:AW303354	-	129.50	286.91	7.3e-07	671	AV651758 AV651758 GLC Homo sap
gb_est45:AW303354	-	129.50	286.85	7.4e-07	675	AW173259 xj85c08.x1 Soares.NFL
gb_est45:AW303354	-	129.50	286.68	7.5e-07	687	AV700236 AV700236 GRC Homo sap
gb_est45:AW303354	-	129.50	286.53	7.7e-07	697	BG562602 602581372F1 NIH_MGC_76
gb_est45:AW303354	-	129.50	284.88	9.5e-07	822	BG564407 602584345F1 NIH_MGC_76
gb_est45:AW303354	-	129.50	284.63	9.8e-07	843	AL563977 AL563977 LIT_NFL001_NH
gb_est45:AW303354	-	129.50	287.69	6.6e-07	552	D83828 HUMHBC4837 Human pancrea
gb_est45:AW303354	-	128.50	291.63	4.0e-07	331	AV657144 AV657144 GLC Homo sap
gb_est45:AW303354	-	128.50	291.04	4.3e-07	351	AV658016 AV658016 GLC Homo sap
gb_est45:AW303354	-	128.50	290.79	4.4e-07	360	AV659081 AV659081 GLC Homo sap
gb_est45:AW303354	-	128.50	290.73	4.5e-07	362	AV682237 AV682237 GXB Homo sap
gb_est45:AW303354	-	128.50	289.81	5.0e-07	397	AV654397 AV654397 GLC Homo sap
gb_est45:AW303354	-	128.50	289.51	5.2e-07	409	AV645675 AV645675 GLC Homo sap
gb_est45:AW303354	-	128.50	289.08	5.5e-07	427	H93888 yv08g08.r1 Soares fetal
gb_est45:AW303354	-	128.50	287.11	7.1e-07	520	AV720273 AV720273 GLC Homo sap
gb_est45:AW303354	-	128.50	285.56	8.7e-07	606	AV683948 AV683948 GRC Homo sap
gb_est45:AW303354	-	128.50	285.56	8.7e-07	607	AV654309 AV654309 GLC Homo sap
gb_est45:AW303354	-	128.50	285.49	8.8e-07	611	AV655831 AV655831 GLC Homo sap
gb_est45:AW303354	-	128.50	284.75	9.6e-07	658	AV645552 AV645552 GLC Homo sap
gb_est45:AW303354	-	128.50	283.37	1.2e-06	755	AV683159 AV683159 GRC Homo sap
gb_est45:AW303354	-	128.00	288.51	6.0e-07	402	BF72811 yu07f10.r1 Soares fetal
gb_est45:AW303354	-	128.00	287.40	6.9e-07	449	BF990562 RC6-GN0071-271000-013
gb_est45:AW303354	-	127.00	283.65	1.1e-06	653	AW112118 MT15218 mouse liver, d
gb_est45:AW303354	-	127.00	283.85	1.1e-06	506	A1065103 HA0939 Human fetal liv
gb_est45:AW303354	-	127.00	283.48	1.1e-06	525	A1333042 qq3ln04.x1 Soares.Nhm
gb_est45:AW303354	-	127.00	282.43	1.3e-06	583	A1336247 qt45b07.x1 Soares.feta
gb_est45:AW303354	-	127.00	281.38	1.5e-06	647	AV690405 AV690405 GRC Homo sap

gb_est29:AL551326 + 127.00 277.49 2.4e-06 955 ! AL551326 AL551326 LTI_NFL006
gb_est42:AW111217 - 126.50 281.69 1.4e-06 558 ! AW111217 MT2416 mouse liver,
gb_est95:BG008148 + 126.50 281.21 1.5e-06 585 ! BG008148 OV4-GN0314-291100-6
gb_est17:AT246161 - 126.00 281.87 1.4e-06 487 ! AT246161 qi29B06.x1 Soares_N
seq_name: gb_est45:AW303354
seq_documentation_block:
LOCUS AW303354 443 bp mRNA EST 18-JAN-2000
DEFINITION xv18p01.x1 Soares.NFL.T_GBC.s1 Homo sapiens cDNA clone
IMAGE:2813449 3' similar to SW:IBP_CARCR P00993 CHELONIANIN ; , mRNA
sequence.
ACCESSION AW303354
VERSION AW303354.1 GI:6713034
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco.
FEATURES
source
1..443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2813449"
/lab_host="Soares.NFL.T_GBC.s1"
/notes="Organ: pooled; Vector: pT7T3D-pac (Pharmacia) with
a modified polylinker; Site.1: Not I; Site.2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI_CGAP_CCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 137 a 84 c 77 g 145 t
ORIGIN

alignment_scores:
Quality: 433.00 Length: 78
Ratio: 5.551 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-177 x AW303354/rev ..
Align seg 1/1 to reverse of: AW303354 from: 1 to: 443
22 LeuLeuGlyGlyValAsnLysIleAlaGluLysIleCysGlyAspLeu 38
|||||
441 TTATTGGGTGGTGAATAAAATTCGGGAGAGATATGGAGACTCAA 392
38 saspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
|||||
391 AGATCCCTGCAATTTGGACATGAATTTGGGAAGCTGCTATGAAGTTCACT 342
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
|||||
341 TTAGATATTCTTACACAGACCTCCAAAAGATGTGAAACTTTTGTCTTC 292

224 GATTTTCTGTTCACTGGCTGCAATGGCAACCTTAATAGCTTCAGAGCTG 275

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
Fax: 81-45-503-9216
224

84 ysileGluArgGluValAlaCysValAlaLysTyrLysProPro 98
 |||||::: ::|||:::||||
 274 AAATAGACTGTGATGCTACTGTCAGAAACATACAAAACCT 317

seq_name: gb_gss32:AZ763929

seq_documentation_block: 570 bp DNA GSS 16-FEB-2001.
 LOCUS AZ763929 IM0559C11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0559C11 R, DNA sequence.

ACCESSION AZ763929
 VERSION AZ763929.1 GI:12878349
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 570)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D., Weis,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00
 Plate: 0559 row: C column: 11

Seq primer: CACACAGAAACAGCATGACC

Class: plasmid ends

High quality sequence stop: 570.

FEATURES
 source

1. 570
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0559C11"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gil4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 179 a 123 c 126 g 142 t

ORIGIN

alignment_scores:
 Quality: 235.00 Length: 60

Ratio: 4.519 Gaps: 0
 Percent Similarity: 86.667 Percent Identity: 65.000

alignment_block:

US-09-215-435-177 x AZ763929/rev ..

Align seg 1/1 to reverse of: AZ763929 from: 1 to: 570

39 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisph 55
 ||||| ||||| ||||| ::::: ||||| ||||| ||||| ||||| |||||
 208 GATCCCTGCTGTTGGATGCTGGACCTGGCAGCTGCTATGAAGTACATT 159
 55 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 72
 ||||| ||||| ||||| ::::: ||||| ||||| ||||| ||||| |||||
 158 TAGATTTTCTACAAACCAACTGCCAAACAGCTGCAGATTTTCTGTTC 109
 72 erGlyCysAsnGlyAsnLeuAsnAsnAsnPheLysLeuLysIleGluArgGlu 88
 ::||| ||||| ||||| ::||| ||||| ||||| ||||| ||||| |||||
 108 CTGGCTGCAATGGCAACTTTAATACTTCAAGCTGAAATAGACTGTGAT 59
 89 ValAlaCysValAlaLysTyrLysProPro 98
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 58 GTAACTTGTCTCAGAAACATACAAAACCCCT 29

seq_name: gb_est18:AI256949

seq_documentation_block:

LOCUS AI256949 552 bp mRNA EST 17-NOV-1998
 DEFINITION LP04037.5prime LP Drosophila melanogaster larval-early pupal pot2
 Drosophila melanogaster cDNA clone LP04037 5prime, mRNA sequence.

ACCESSION AI256949

VERSION AI256949.1 GI:3864474

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM

Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 552)
 AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
 P., Lewis,S. and Rubin,G.M.

TITLE BDGP/HHMI Drosophila EST Project

JOURNAL Unpublished (1997)

COMMENT Contact: Harvey, D.

G. M. Rubin-Molecular and Cell Biology

University of California Berkeley

539 USA, Berkeley, CA 94720-3200, USA

Fax: 510 643 9947

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu

Sequence is the complete cDNA Insert. hit genomic sequence DS00216;
 hit genomic sequence DS02911; hit genomic sequence DS03364

Plate: 40 row: D column: 1

High quality sequence stop: 552.

FEATURES

source

1. 552
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="LP04037"
 /clone_lib="LP Drosophila melanogaster larval-early pupal
 pot2"
 /sex="male and female"
 /dev_stage="larvae-pupae"
 /lab_host="DH5-alpha"
 /note="Organ: whole body; Vector: pot2; Site_1: EcoRI;
 Site_2: XhoI; Sized fractionated cDNAs were directly
 ligated into pot2. Plasmid cDNA library."

BASE COUNT 154 a 124 c 122 g 152 t

ORIGIN

alignment_scores:
 Quality: 132.50 Length: 101

(total) fetus material with a Not I - oligo(dT) primer [5
TCTTACCAATCGAAGTGGGAGCGCGCTAATTTTCTTTTCTTTT
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo.

```

Ratio: 2.366          Gaps: 4
Percent Similarity: 55.446  Percent Identity: 31.683

alignent_block:
US-09-215-435-177 x AI256949 ..

Align seg 1/1 to: AI256949 from: 1 to: 552

12 PhePheIlePheCysSerLeuAsnThrLeuLeuLeuGlyGlyVal..... 26
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
31 TTCTAGTATTGCGGCGTGT...CTGCTCTCGCAGCCGGAATTCGTGC 77
27 .....AsnLysIleAlaGluLys.....IleCysG 35
|||:|||||:|||||:|||||:|||||:|||||:|||||:
78 ACCTCCCTCGGATGTCGTCGTGGAGAAGCGGAGGAGCCGCGGTGGCA 127
|||:|||||:|||||:|||||:|||||:|||||:|||||:
35 LysAspLeuLysAsp.....ProCysLysLeu 43
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 AACCCCTCGCTGATCCAGTGCAGTCATCTTCCGAGGACTGTCATCAG 177
44 AspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheTyrAs 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 CCCAAGAAACCGGTGCGTGTTCGTCGTCTGTCATCGCTACGCCCTACAA 227
nArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyA 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 CGTGGATACACATCCTCGCAGGAGTTCGTTTACGGTGGATGTGCCGCA 277
77 snLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCysValAla 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 ACAAGAACAACTTCGAATCCAGAGGACGTCGCGAACACGCGATGTTGGTA 327
94 Lys 94
|||
328 AAG 330

```

seq_documentation_block:	27-AUG-1998	EST	02-MAR-1995
LOCUS	AI015730	mRNA	mRNA
DEFINITION	ot7lb08.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:1622199 3' similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN);		
); mRNA sequence.		
ACCESSION	AI015730		
VERSION	AI015730.1	GI:3230066	
KEYWORDS	EST		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 496)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov This clone is available royalty-free through LLNL : contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 2157 Std Error: 0.00 Seq primer: -40m13 fwd. Et from Amersham High quality sequence stop: 272. Location/Qualifiers		
FEATURES			
source	1..496		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1622199"		
	/clone_lib="Soares_total_fetus_Nb2HF8_9w"		
	/dev_stage="8-9 weeks"		
	/lab_host="PH10B"		
	/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA		

```
High quality sequence stop: 281.
FEATURES
  source
    1. .350
      /db_xref="taxon:501869"
      /db_xref="taxon:9606"
      /clone_lib="IMAGE:84812"
      /clone_lib="Stratagene liver (#937224)"
      /sex="male"
      /dev_stage="49 years old"
      /lab_host="SOLR cells (kanamycin resistant)"
      /note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dr. Hepatectomy from normal male caucasian. Average insert
size: 1.1 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
CTCGAGTCTTTTTTTTTTTT 3'."
BASE COUNT      78 a   92 c  103 g   76 t      1 others
ORIGIN
alignment_scores:
  Quality: 129.50      Length: 70
  Ratio: 3.083        Gaps: 2
  Percent Similarity: 60.000      Percent Identity: 41.429
alignment_block:
US-09-215-435-177 x T74655 ..
Align seg 1/1 to: T74655 from: 1 to: 350
24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||:||||:||||:||||:||||:||||:||||:||||:||||:
2 GGTGGCAACTGGTAACTGAAGTCACCAAGAAA.....GA 36
38 sAspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
||||| |||:||||:||||:||||:||||:||||:||||:||||:
37 AGATTCTCCAGCTGGGTACTCGCGCGGCTCCTGCATGGGAATGACCA 86
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||:||||:||||:||||:||||:||||:||||:||||:
87 GCAGGTATTTCTATATGATGATCATCCATGGCTGTGAGACTTTCCAGTAC 136
72 SerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgG1 88
||||| |||:||||:||||:||||:||||:||||:||||:||||:
137 GCGGCTGTCATGGCAACGTCGTAACACTTCGTCACAGAAAGAGTGCT 186
88 uValAlaCys 91
|||||
187 GCAGACCTGC 196
seq_name: gb_est7:AA424642
seq_documentation_block:
LOCUS      AA424642      413 bp      mRNA      EST      16-OCT-1997
DEFINITION      zv91d02.r1 Soares.NhMPu.S1 Homo sapiens cDNA clone IMAGE:767139 5'
similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN);, mRNA sequence.
ACCESSION      AA424642
VERSION      AA424642.1 GI:2103595
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 413)
AUTHORS      Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
JOURNAL      Unpublished (1997)
COMMENT      Contact: Wilson RK
Washington University School of Medicine
```

```
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham.
FEATURES
  Location/Qualifiers
    1. .413
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="IMAGE:767139"
      /clone_lib="Soares.NhMPu.S1"
      /tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
      /lab_host="DH10B"
      /note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT      93 a  114 c  126 g   80 t
ORIGIN
alignment_scores:
  Quality: 129.50      Length: 70
  Ratio: 3.083        Gaps: 2
  Percent Similarity: 60.000      Percent Identity: 41.429
alignment_block:
US-09-215-435-177 x AA424642 ..
Align seg 1/1 to: AA424642 from: 1 to: 413
24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||:||||:||||:||||:||||:||||:||||:||||:
88 GGTGGCAACTGGTAACTGAAGTCACCAAGAAA.....GA 122
38 sAspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
||||| |||:||||:||||:||||:||||:||||:||||:||||:
123 AGATTCTCCAGCTGGGTACTCGCGCGGCTCCTGCATGGGAATGACCA 172
55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||:||||:||||:||||:||||:||||:||||:||||:
173 GCAGGTATTTCTATATGATGATCATCCATGGCTGTGAGACTTTCCAGTAC 222
72 SerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgG1 88
||||| |||:||||:||||:||||:||||:||||:||||:||||:
223 GCGGCTGTCATGGCAACGTCGTAACACTTCGTCACAGAAAGAGTGCT 272
88 uValAlaCys 91
|||||
273 GCAGACCTGC 282
seq_name: gb_est14:AF063548
seq_documentation_block:
LOCUS      AF063548      511 bp      mRNA      EST      16-JUN-1999
DEFINITION      AF063548 Homo sapiens library (Yu Y) Homo sapiens cDNA clone HA0104
, mRNA sequence.
ACCESSION      AF063548
VERSION      AF063548.1 GI:5080918
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```
/clone_lib="Human fetal liver cdna library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/p3"
/note="vector: pCDNA1"
BASE COUNT      116 a   142 c   157 g   106 t
ORIGIN
```

```
alignment_scores:
  Quality: 129.50      Length: 70
  Ratio: 3.083        Gaps: 2
  Percent Similarity: 60.000  Percent Identity: 41.429
alignment_block:
US-09-215-435-177 x AI207711 ..
Align seg 1/1 to: AI207711 from: 1 to: 521
```

```
24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
|||||  |||:||||:||||:||||:||||:||||:||||:||||:||||:
181 GGTGGCAACTGGTAAGTCAAGTACCAAGAAA.....GA 215

38 sAspProCysLysLeuAspMetAsnPhcPheCysTyrGluValHisP 55
:||||:||||:||||:||||:||||:||||:||||:||||:||||:
216 AGATTCTCCAGCTGGCTACTCGCGCGTCCCTGCATGGGAATGACCA 265

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
|||||  |||:||||:||||:||||:||||:||||:||||:||||:
266 GCAGGTATTCTATAATGTATCATCCATGCCTGTGAGACTTTCCAGTAC 315

72 SerGlyCysAsnGlyAsnLeuAsnPhcLysLeuLysIleGluArgG 88
:|||||  |||:||||:||||:||||:||||:||||:||||:||||:
316 GCGGCTGCTGATGGCAACGGTAACACTTCGTCACAGAAAGGAGTGCT 365

88 uValAlaCys 91
:|||||
366 GCAGACCTGC 375
```

seq_name: gb_est17:AI207590

```
seq_documentation_block:
LOCUS      AI207590      571 bp      mRNA      EST      11-NOV-1999
DEFINITION HA2954 Human fetal liver cdna library Homo sapiens CDNA, mRNA
sequence.
ACCESSION  AI207590
VERSION    AI207590.1 GI:6361600
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 571)
AUTHORS   Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
TITLE     Expression profile analysis of a human fetal liver cdna library
JOURNAL   Unpublished (1998)
COMMENT   Contact: Chenggang Zhang
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Email: zhang_chenggang@hotmail.com.
FEATURES   Location/Qualifiers
            source
            1..571
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="Human fetal liver cdna library"
            /tissue_type="liver"
            /dev_stage="fetal"
            /lab_host="MC1061/p3"
            /note="vector: pCDNA1"
```

```
BASE COUNT      113 a   174 c   149 g   135 t
ORIGIN
```

alignment_scores:

```
  Quality: 129.50      Length: 70
  Ratio: 3.083        Gaps: 2
  Percent Similarity: 60.000  Percent Identity: 41.429
```

alignment_block:
US-09-215-435-177 x AI207590/rev ..

Align seg 1/1 to reverse of: AI207590 from: 1 to: 571

```
24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
|||||  |||:||||:||||:||||:||||:||||:||||:||||:
523 GGTGGCAACTGGTAAGTCAAGTACCAAGAAA.....GA 489

38 sAspProCysLysLeuAspMetAsnPhcPheCysTyrGluValHisP 55
:||||:||||:||||:||||:||||:||||:||||:||||:||||:
488 AGATTCTCCAGCTGGCTACTCGCGCGTCCCTGCATGGGAATGACCA 439

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
|||||  |||:||||:||||:||||:||||:||||:||||:||||:
438 GCAGGTATTCTATAATGTATCATCCATGCCTGTGAGACTTTCCAGTAC 389

72 SerGlyCysAsnGlyAsnLeuAsnPhcLysLeuLysIleGluArgG 88
:|||||  |||:||||:||||:||||:||||:||||:||||:||||:
388 GCGGCTGCTGATGGCAACGGTAACACTTCGTCACAGAAAGGAGTGCT 339

88 uValAlaCys 91
:|||||
338 GCAGACCTGC 329
```

seq_name: gb_est40:AV658823

```
seq_documentation_block:
LOCUS      AV658823      626 bp      mRNA      EST      07-SEP-2000
DEFINITION AV658823 GLC Homo sapiens cdna clone GLCFQE08 3', mRNA sequence.
ACCESSION  AV658823
VERSION    AV658823.1 GI:9879837
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 626)
AUTHORS   Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.
, Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE     Homo sapiens cdna clone
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
```

FEATURES
source

```
1..626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCFQE08"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
```

```
BASE COUNT      163 a   157 c   183 g   120 t   3 others
ORIGIN
```

us-09-215-435-177.rst

Thu Nov 15 10:51:54 2001

726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 128 a 195 c 173 g 148 t

ORIGIN

alignment_scores:
Quality: 129.50 Length: 70
Ratio: 3.083 Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x AI807643/rev ..

Align seg 1/1 to reverse of: AI807643 from: 1 to: 644

24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||.....:|||||
86 GTGGGCAACTGTAAGTACCAAGAAA.....GA 120

38 saspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
||||| |||.....:|||||
121 AGATTCTCCAGCTGGGTACTCGCGCGGTCCTGCATGGGAATGACCA 170

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||.....:|||||
171 GCAGGTATTTCTATATGTTACATCCATGGCCTGTGAGACTTTTCCAGTAC 220

72 SerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgG1 88
||||| |||.....:|||||
221 GCGCGCTGCATGGCAACGGTAACACTTCGTGCACAGAAAGGAGTGCT 270

88 uValAlaCys 91
:|||||
271 GCAGACCTGC 280

seq_name: gb_est25:AI807643

seq_documentation_block: 644 bp mRNA EST 19-DEC-1999
LOCUS AI807643.1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:2358919 3' similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN
);, mRNA sequence.

ACCESSION AI807643
VERSION AI807643.1 GI:5394209

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 644)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1328 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 462.
Location/Qualifiers
1. .644

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2358919"
/lab_host="DH108"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19w, testis NHR, and B-cell
NCI-CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,

1. .662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/p3"
/note="Vector: pCDNA1"

BASE COUNT 161 a 177 c 182 g 138 t 4 others
ORIGIN

726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 128 a 195 c 173 g 148 t

ORIGIN

alignment_scores:
Quality: 129.50 Length: 70
Ratio: 3.083 Gaps: 2
Percent Similarity: 60.000 Percent Identity: 41.429

alignment_block:
US-09-215-435-177 x AI807643/rev ..

Align seg 1/1 to reverse of: AI807643 from: 1 to: 644

24 GlyGly.....ValAsnLysIleAlaGluLysIleCysGlyAspLeuLy 38
||||| |||.....:|||||
520 GTGGGCAACTGTAAGTACCAAGAAA.....GA 486

38 saspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisP 55
||||| |||.....:|||||
485 AGATTCTCCAGCTGGGTACTCGCGCGGTCCTGCATGGGAATGACCA 436

55 heArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPhe 71
||||| |||.....:|||||
435 GCAGGTATTTCTATATGTTACATCCATGGCCTGTGAGACTTTTCCAGTAC 386

72 SerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgG1 88
||||| |||.....:|||||
385 GCGCGCTGCATGGCAACGGTAACACTTCGTGCACAGAAAGGAGTGCT 336

88 uValAlaCys 91
:|||||
335 GCAGACCTGC 326

seq_name: gb_est15:AI065076

seq_documentation_block: 662 bp mRNA EST 11-NOV-1995
LOCUS AI065076 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
DEFINITION HA0910 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.

ACCESSION AI065076
VERSION AI065076.1 GI:6359348

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 662)

AUTHORS Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.

Expression profile analysis of a human fetal liver cDNA library
Unpublished (1998)

JOURNAL

COMMENT

Contact: Yongtao Yu

Department of Hematology

Beijing Institute of Radiation Medicine

27 Taiping Road, Beijing 100850, P.R.China

Tel: 0086-10-68159479

Fax: 0086-10-68214653

Email: yyt48@yahoo.com.

Location/Qualifiers

1. .662

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human fetal liver cDNA library"

/tissue_type="liver"

/dev_stage="fetal"

/lab_host="MC1061/p3"

/note="Vector: pCDNA1"

BASE COUNT 161 a 177 c 182 g 138 t 4 others

ORIGIN

```
alignment_scores:
    Quality: 129.50      Length: 70
    Ratio:   3.083       Gaps: 2
    Percent Similarity: 60.000     Percent Identity: 41.429

alignment_block:
US-09-2I5-435-177 x AI065076 ..

Align seg 1/1 to: AI065076 from: 1 to: 662

        24 GlyGly.....ValAsnLysIlleAlaGluLysIlleCysGlyAspLeuLy 38
           |||||         |:::|||||
        368 GGTTGGCAACTGGTAACTACCAAGAACAAAAA.....GA 402
                               :::

        38 sAspProCysLysLeuAspMetAsnPheGlySerCysTyrgluValHisP 55
           |||||         |:|:|||
        403 AGATTCTGCAGCTGGGTACTCGGCCGGTCCTGCATGGGAATGACCA 452

        55 heArgtyrPheTyraSnArgrThrSerLysArgCysgluthrPheValPhe 71
           |||||         |::::|
        453 GCAGGTATTCTATAATGGTGACATCCATGGCCTGTGAGACTTTCACAGTAC 502
                               :|||

        72 SerGlyCysAsnGlyAsnLeuAsnPheLysLysLysIlleGluargl 88
           :::|||||         |::::|
        503 GCGCGTGCATGGCCAACGGTAACAACATTCGTGCACAGAAAAGGAGTGCT 552
                               :|||

        88 uValalaCys 91
           :|||

        553 GCAGAGCTGC 562
```

us-09-215-435-177.rst

Thu Nov 15 10:51:54 2001

OM of: US-09-215-435-177_COPY1_24 to: GenEmbl:* out_format : pfs

Date: Nov 15, 2001 12:40 AM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -FGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435 -CGN1_1_11006 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY1_24

Query length: 24

Database: GenEmbl:*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 7698.940000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
gb_pr9:HSDJ60101	+ 120.00	266.48	1.3e-06	33458	AL109656 Human DNA sequence f
gb_pr9:AC023971	+ 120.00	253.75	6.8e-06	153788	AC023971 Homo sapiens chromo
gb_hlg10:AC024934	- 67.00	117.72	255.00	209777	AC024934 Homo sapiens chromo
gb_hlg3:AC012143	+ 66.00	117.32	271.32	164426	AC012143 Homo sapiens clone
gb_pr2:AC008275	+ 66.00	117.13	274.99	166504	AC008275 Homo sapiens BAC cl
gb_hlg22:AL451047	+ 62.00	107.31	968.73	161545	AL451047 Homo sapiens chromo
gb_hlg13:AC055798	- 62.00	107.02	1.0e+03	167314	AC055798 Homo sapiens chromo
gb_hlg4:AC014616	+ 60.00	104.74	1.3e+03	120188	AC014616 Drosophila melanoga
gb_hlg2:AC010577	+ 60.00	103.98	1.5e+03	131684	AC010577 Drosophila melanoga
gb_hlg10:AC025616	- 60.00	103.61	1.6e+03	137670	AC025616 Homo sapiens clone
gb_in2:AE003846	+ 60.00	97.15	3.6e+03	298469	AE003846 Drosophila melanoga
gb_hlg7:AC021204	+ 59.50	101.18	2.1e+03	158357	AC021204 Homo sapiens chromo
gb_hlg8:AC022211	+ 59.00	98.35	3.1e+03	191151	AC022211 Homo sapiens chromo
gb_pr1:AC005002	- 58.00	99.73	2.6e+03	119841	AC005002 Homo sapiens clone
gb_pr9:HSJ612B15	- 58.00	99.19	2.7e+03	127820	AL049597 Human DNA sequence
gb_hlg21:AL355981	+ 58.00	97.54	3.4e+03	155871	AL355981 Homo sapiens chromo
gb_hlg1:AC004688	+ 58.00	96.82	3.7e+03	169794	AC004688 Plasmodium falcipar
gb_hlg1:AC004709	+ 58.00	95.62	4.3e+03	196149	AC004709 Plasmodium falcipar
gb_in3:CEF5A11	- 57.50	110.73	624.36	27598	Z72511 Caenorhabditis elegans
gb_r01:AF084363	- 57.00	98.02	3.2e+03	108752	AF084363 Mus musculus D6M5e
gb_hlg18:AC087882	+ 57.00	96.19	4.0e+03	135445	AC087882 Homo sapiens chromo
gb_hlg19:AF282854	+ 57.00	96.10	4.1e+03	137020	AF282854 Homo sapiens chromo
gb_r01:AC003061	+ 57.00	94.66	4.9e+03	162691	AC003061 Mouse Chromosome 6
gb_pr2:AC007876	+ 57.00	94.55	5.0e+03	164879	AC007876 Homo sapiens BAC cl
gb_hlg10:AC024931	+ 57.00	93.79	5.5e+03	180519	AC024931 Homo sapiens chromo
gb_pr3:AC017082	+ 57.00	93.64	5.6e+03	183783	AC017082 Homo sapiens clone
gb_hlg5:AC016679	+ 57.00	93.47	5.7e+03	187695	AC016679 Homo sapiens chromo
gb_hlg1:AC007305	+ 57.00	92.25	6.7e+03	217115	AC007305 Mus musculus chromo
gb_r01:CN007004	- 56.00	134.49	29.68	1021	AL423434 T3 end of clone AZ0AA
gb_hlg3:AC011770	+ 56.00	100.93	2.2e+03	56774	AC011770 Homo sapiens chromos
gb_hlg18:AC090257	+ 56.00	99.36	2.7e+03	68585	AC090257 Homo sapiens chromos
gb_pr1:AC004024	- 56.00	98.16	3.1e+03	79164	AC004024 Homo sapiens 12q24.2
gb_hlg11:AC027448	- 56.00	98.11	3.2e+03	79595	AC027448 Homo sapiens chromos
gb_in1:AC024854	+ 56.00	97.73	3.3e+03	83297	AC024854 Caenorhabditis elega
gb_in3:CELY71F9B	+ 56.00	97.73	3.3e+03	83297	AC024201 Caenorhabditis elega
gb_hlg1:AC008947	+ 56.00	97.19	3.5e+03	88948	AC008947 Homo sapiens chromos
gb_pr3:AC008860	- 56.00	95.73	4.3e+03	108886	AC008860 Homo sapiens chromo
gb_hlg10:AC025444	- 56.00	95.53	4.4e+03	108405	AC025444 Homo sapiens chromo
gb_hlg2:AC010044	- 56.00	94.70	4.9e+03	119844	AC010044 Drosophila melanoga
gb_p11:AB001684	- 56.00	92.79	6.2e+03	150613	AB001684 Chlorocella vulgaris
gb_hlg23:AL590558	+ 56.00	92.59	6.4e+03	154269	AL590558 Homo sapiens chromo

gb_hlg14:AC068206 + 56.00 92.11 6.8e+03 163260 ! AC068206 Homo sapiens chr
gb_hlg11:AC026360 + 56.00 91.92 7.0e+03 167206 ! AC026360 Homo sapiens chr
gb_hlg7:AC020687 - 56.00 91.79 7.1e+03 169635 ! AC020687 Homo sapiens chr
gb_hlg19:AL133513 - 56.00 91.62 7.2e+03 173218 ! AL133513 Homo sapiens chr

seq_name: gb_pr9:HSDJ60101

seq_documentation_block:

LOCUS HSDJ60101 33458 bp DNA PRI 04-APR-2001
DEFINITION Human DNA sequence from clone RP4-60101 on chromosome 20. Contains
the gene for a novel protein with Kunitz/Bovine pancreatic trypsin
inhibitor domain, the 5' end of a novel gene, part of the gene for
a novel protein with WAP-type (Whey Acidic Protein) 'four-disulfide
core' domain, ESTs and GSSs, complete sequence.
ACCESSION AL109656
VERSION AL109656.10 GI:10944958
KEYWORDS HTG; Kunitz; WAP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Laird.G.
1 (bases 1 to 33458)
Direct Submission
Submitted (03-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On Oct 21, 2000 this sequence version replaced gi:9795206.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RP4-60101. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP3-447F3 is at 33359 in this sequence.
The true right end of clone RP4-688G8 is at 100 in this sequence.
RP4-60101 is from the library RPCI-4 constructed by the group of
Pietter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest.

location/Qualifiers

1. .33458

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="20"

/clone="RP4-60101"

/clone_lib="RPCI-4"

3. .225

repeat_region

/note="ILM4 repeat: matches 3977. .4203 of consensus"

225. .428

repeat_region

/note="ILM4 repeat: matches 4321. .4525 of consensus"

429. .696

repeat_region

/note="ALuJb repeat: matches 2. .302 of consensus"

697. .778

FEATURES

source

[illegible]

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
 |||||
 25414 ATGAAGTCTGCCAGCTGGGATTTCTTCTAAGATTCTTCATCTTCGTC 25463

17 rleuAsnThrLeuLeuGly 24
 |||||

25464 ATTGAATACCTGTATGGGT 25485

seq_name: gb_htg9:AC023971

seq_documentation_block:

LOCUS AC023971 153788 bp DNA HTG 07-APR-2000
 DEFINITION Homo sapiens chromosome 20 clone RP11-770N8 map 20, WORKING DRAFT
 SEQUENCE, 13 unordered pieces.

ACCESSION AC023971

VERSION AC023971.2 GI:7523758

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 153788)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 20, clone RP11-770N8

REFERENCE 2 (bases 1 to 153788)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
 Boukhalter,B., Brown,A., Burkett,G., Campoliano,A., Castie,A.,
 Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
 DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
 Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Landers,T., Larcocque,K., Lehoczy,J., Levine,R.,
 Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
 McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M.,
 Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
 Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
 Travers,M., Triggilio,J., Vassiliev,H., Viel,R., Vo.A., Willson,B.,
 Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
 Zody,M.

TITLE
 JOURNAL

COMMENT

Direct Submission
 Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 7, 2000 this sequence version replaced gi:7008875.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6573

Center clone name: 770_N_8

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 145038 bases at least Q40

Consensus quality: 148490 bases at least Q30

Consensus quality: 150621 bases at least Q20

Insert size: 150000; agarose-fp

Insert size: 152588; sum-of-ctnigs

Quality coverage: 5.2 in Q20 bases; agarose-fp

Quality coverage: 5.1 in Q20 bases; sum-of-ctnigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1333: contig of 1333 bp in length
 * 1334 1433: gap of 100 bp
 * 1434 3969: contig of 2536 bp in length
 * 3970 4069: gap of 100 bp
 * 4070 7211: contig of 3142 bp in length
 * 7212 7311: gap of 100 bp
 * 7312 13518: contig of 6207 bp in length
 * 13519 13618: gap of 100 bp
 * 13619 20876: contig of 7258 bp in length
 * 20877 20976: gap of 100 bp
 * 20977 28838: contig of 7862 bp in length
 * 28839 28938: gap of 100 bp
 * 28939 37590: contig of 8652 bp in length
 * 37591 37690: gap of 100 bp
 * 37691 45814: contig of 8124 bp in length
 * 45815 45914: gap of 100 bp
 * 45915 62601: contig of 16687 bp in length
 * 62602 62701: gap of 100 bp
 * 62702 81817: contig of 19116 bp in length
 * 81818 81917: gap of 100 bp
 * 81918 100273: contig of 18356 bp in length
 * 100274 100373: gap of 100 bp
 * 100374 125266: contig of 24893 bp in length
 * 125267 125366: gap of 100 bp
 * 125367 153788: contig of 28422 bp in length.

FEATURES

Source

1. 153788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20"
 /clone="RP11-770N8"
 /clone_lib="RPCI-11 Human Male BAC"
 misc_feature 1. 1333
 /note="assembly_fragment"
 misc_feature 1434..3969
 /note="assembly_fragment"
 misc_feature 4070..7211
 /note="assembly_fragment"
 misc_feature 7312..13518
 /note="assembly_fragment"
 misc_feature 13619..20876
 /note="assembly_fragment"
 misc_feature 20977..28838
 /note="assembly_fragment"
 misc_feature 28939..37590
 /note="assembly_fragment"
 misc_feature 37691..45814
 /note="assembly_fragment"
 misc_feature 45915..62601
 /note="assembly_fragment"
 misc_feature 62702..81817
 /note="assembly_fragment"
 misc_feature 81918..100273
 /note="assembly_fragment"
 misc_feature 100374..125266
 /note="assembly_fragment"
 misc_feature 125367..153788
 /note="assembly_fragment"

BASE COUNT 44583 a 30904 c 30568 g 46532 t 1201 others

Thu Nov 15 10:51:55 2001

```

ORIGIN
alignment_scores:
  Quality: 120.00      Length: 24
  Ratio: 5.000         Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_1_24 x AC023971/rev ..
Align seg 1/1 to reverse of: AC023971 from: 1 to: 153788

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSse 17
|||||
135635 ATGAGTCTGCCAAGCTGGGATTTCTTACAGATCTTCATCTTCGTC 135586

17 rLeuAsnThrLeuLeuGly 24
|||||
135585 ATTGATACCTGTTATTGGGT 135564

seq_name: gb_hcg10:AC024934

seq_documentation_block:
LOCUS AC024934 209777 bp DNA HTG 30-DEC-2000
DEFINITION Homo sapiens chromosome 12 clone RP11-329A19, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
ACCESSION AC024934
VERSION AC024934.29 GI:11955487
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 209777)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunga-H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K.K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,Z., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newtonson,J., Newtonson,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogutu,M., Okwundu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williams,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.

```

```

and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 209777)
REFERENCE
Worley, K.C.
Direct Submission
Submitted (03-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 29, 2000 this sequence version replaced gi:11192044.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project: Information
Center project name: HAJA
Center clone name: RP11-329A19
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 197273 bases at least Q40
Consensus quality: 202546 bases at least Q30
Estimated insert size: 205790 bases at least Q20
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 5.5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 33513: contig of 33513 bp in length
* 33514: gap of unknown length
* 33614: contig of 24592 bp in length
* 58205: gap of unknown length
* 58206: contig of 3310 bp in length
* 58206: gap of unknown length
* 91615: gap of unknown length
* 91616: contig of 22902 bp in length
* 114617: gap of unknown length
* 114618: contig of 18143 bp in length
* 132861: gap of unknown length
* 132861: contig of 19279 bp in length
* 152239: gap of unknown length
* 152240: contig of 17505 bp in length
* 152340: gap of unknown length
* 169845: contig of 15307 bp in length
* 169945: gap of unknown length
* 185252: contig of 9207 bp in length
* 185352: gap of unknown length
* 194559: contig of 9422 bp in length
* 194658: gap of unknown length
* 204081: contig of 1949 bp in length
* 204180: gap of unknown length
* 204181: contig of 1101 bp in length
* 206130: gap of unknown length
* 206223: contig of 1244 bp in length
* 207331: gap of unknown length
* 207431: contig of 1244 bp in length
* 208675: gap of unknown length
* 208775: contig of 1003 bp in length.
Location/Qualifiers
1. 209777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-329A19"

```

FEATURES
source

BASE COUNT 65720 a 39190 c 38969 g 64529 t 1369 others
ORIGIN

alignment_scores:
Quality: 67.00 Length: 22
Ratio: 4.188 Gaps: 1
Percent Similarity: 72.727 Percent Identity: 68.182

alignment_block:

US-09-215-435-177_COPY_1_24 x AC024934/rev ..

Align seg 1/1 to reverse of: AC024934 from: 1 to: 209777

7 GlyPheLeuLeuArgPhePhe.....IlePheCysSerLeuAs 19

|||||TTTTCTTTAAGATTTTTCCTTATCTTTGATTTTCTGCAGTTGAA 45201

19 nThrLeuLeuLeuGly 24

|||||
45200 TACAATTTGCCAGGT 45185

seq_name: gb_htg3:AC012143

seq_documentation_block:

LOCUS AC012143 164426 bp DNA HTG 10-SEP-2000
DEFINITION Homo sapiens clone RP11-11G20, WORKING DRAFT SEQUENCE, 8 unordered
pieces.

ACCESSION AC012143

VERSION AC012143.3 GI:10047745

KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 164426)

Barren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens, clone RP11-11G20

Unpublished

2 (bases 1 to 164426)

Barren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Farrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,W.

Direct Submission

Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 10, 2000 this sequence version replaced gi:7230063.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WTB8

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3159

Center clone name: 11.G.20

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 158797 bases at least Q40
Consensus quality: 161636 bases at least Q30
Consensus quality: 162651 bases at least Q20
Insert size: 163000; agarose-fp
Quality coverage: 5.7 in Q20 bases; agarose-fp.
Quality coverage: 5.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 33254: contig of 33254 bp in length
* 33255 33354: gap of 100 bp
* 33355 37499: contig of 4145 bp in length
* 37500 37599: gap of 100 bp
* 37600 40875: contig of 3276 bp in length
* 40876 40975: gap of 100 bp
* 40976 91250: contig of 50275 bp in length
* 91251 91350: gap of 100 bp
* 91351 107305: contig of 15955 bp in length
* 107306 107405: gap of 100 bp
* 107406 124992: contig of 17587 bp in length
* 124993 125092: gap of 100 bp
* 125093 147928: contig of 22836 bp in length
* 147929 148028: gap of 100 bp
* 148029 164426: contig of 16398 bp in length.

FEATURES

Location/Qualifiers

1..164426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-11G20"
/clone_lib="RPC1-11 Human Male BAC"
1..33254
/note="assembly_fragment
clone_end:SP6
vector_side:left"
33355..37499
/note="assembly_fragment"
37600..40875
/note="assembly_fragment"
40976..91250
/note="assembly_fragment"
91351..107305
/note="assembly_fragment"
107406..124992
/note="assembly_fragment"
125093..147928
/note="assembly_fragment"
148029..164426
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 52060 a 30107 c 29871 g 51686 t 702 others
ORIGIN

alignment_scores:

Quality: 66.00 Length: 23
Ratio: 3.143 Gaps: 0
Percent Similarity: 91.304 Percent Identity: 52.174

alignment_block:

US-09-215-435-177_COPY_1_24 x AC012143/rev ..

Align seg 1/1 to reverse of: AC012143 from: 1 to: 164426

1 MetIysSerAlaIysLeuGlyPheLeuArgPhePheIlePheCysSe 17
|||||
|||||

Thu Nov 15 10:51:55 2001

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)

VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-465G21. Actual start of this clone is at base position 1 of RP11-294I11; actual end is at base position 166504 of RP11-294I11.

FEATURES	source
Location/Qualifiers	1..166504
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-294I11"
	/clone_lib="RPCI-11"
	1051..1346
repeat_region	/rpt_family="Alu"
	1375..1538
repeat_region	/rpt_family="L1"
	2277..2522
repeat_region	/rpt_family="Alu"
	2523..2580
repeat_region	/rpt_family="(TAAA)n"
	2603..3434
repeat_region	/rpt_family="L1"
	3504..3955
repeat_region	/rpt_family="L1"
	3956..4238
repeat_region	/rpt_family="Alu"
	4239..4311
repeat_region	/rpt_family="L1"
	4514..4595
repeat_region	/rpt_family="(TATG)n"
	4736..4759
repeat_region	/rpt_family="AT-rich"
	5357..5645
repeat_region	/rpt_family="MER4-group"
	5647..5941
repeat_region	/rpt_family="L1"
	5977..5999
repeat_region	/rpt_family="(TA)n"
	6012..6289
repeat_region	/rpt_family="MER4-group"
	6793..12276
repeat_region	/rpt_family="L1"
	15813..16189
repeat_region	/rpt_family="MaLR"
	16572..16605
repeat_region	/rpt_family="(CA)n"
	17466..17829
repeat_region	/rpt_family="MaLR"
	18945..18979
repeat_region	/rpt_family="A-rich"
	19114..19222
repeat_region	/rpt_family="MIR"
	19368..19475
repeat_region	/rpt_family="MIR"
	19479..19518
repeat_region	/rpt_family="AT-rich"
	20420..20475
repeat_region	/rpt_family="AT-rich"
	22046..22169
repeat_region	/rpt_family="MIR"
	22214..22246
repeat_region	/rpt_family="(GAAAA)n"
	22722..23007

24972 ATGCAAGTAATCAATTTCTCTCTGTTAAAGATGCTCTATTTCCTC 24923

17 rleuAsnThrLeuLeuLeu 23
|||||:||||:|||||
24922 TCTAATTCCTTGATTCCT 24904

seq_name: gb_pr2:AC008275

seq_documentation_block:
LOCUS AC008275 166504 bp DNA PRI 30-SEP-2000
DEFINITION Homo sapiens BAC clone RP11-294I11 from 2, complete sequence.
ACCESSION AC008275
VERSION AC008275.3 GI:6598922
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 166504)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
2 (bases 1 to 166504)
Hou, S., Stoneking, T., Yookum, M. and Tucci, S.
The sequence of Homo sapiens BAC clone RP11-294I11
Unpublished
3 (bases 1 to 166504)
Waterston, R.H.
Direct Submission
Submitted (31-JUL-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 166504)
Waterston, R.H.
Direct Submission
Submitted (17-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 166504)
Waterston, R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 17, 1999 this sequence version replaced gi:5757528.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
----- Center project name: H_NH0294I11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

/rpt_family="Alu"
24143..24649
/rpt_family="L2"
24650..24945
/rpt_family="Alu"
24946..25143
/rpt_family="L2"
25470..25780
/rpt_family="Alu"
26007..26209
/rpt_family="L2"
26210..26330
/rpt_family="Alu"
26346..26919
/rpt_family="L2"
27078..27249
/rpt_family="MaLR"
27398..27446
/rpt_family="L2"
27711..27752
/rpt_family="GA-rich"
28051..28242
/rpt_family="MIR"
28273..28378
/rpt_family="MIR"
28415..28782
/rpt_family="MaLR"
28783..29074
/rpt_family="MaLR"
29078..29186
/rpt_family="MaLR"
30033..30125
/rpt_family="(TATG)n"
30150..30635
/rpt_family="MaLR"
31305..31436
/rpt_family="L2"
31814..32012
/rpt_family="MaLR"
32053..32073
/rpt_family="AT-rich"
33095..34300
/rpt_family="L1"
34293..34927
/rpt_family="L1"
34928..35684
/rpt_family="Retroviral"
35698..35848
/rpt_family="L1"
36036..36312
/rpt_family="Alu"
36313..36343
/rpt_family="(CAAA)n"
36344..36365
/rpt_family="Alu"
37140..37856
/rpt_family="Other"
38227..38357
/rpt_family="Alu"
38358..38641
/rpt_family="Alu"
38733..38759
/rpt_family="AT-rich"

alignment_scores:
Quality: 66.00 Length: 23
Ratio: 3.143 Gaps: 0
Percent Similarity: 91.304 Percent Identity: 52.174

alignment_block:
US-09-215-435-177_COPY_1_24 AC008275

Align seg 1/1 to: AC008275 from: 1 to: 166504

1 MetLysSerAlaLysLeuGlyPheLeuArgPhePheIlePheCysse 17
||||: |||:::|||||||:::|||||||
51082 ATGCAAGTAATCAAAATTTCTCTCTGTTAAGATGCTCTATTATTTGTC 51131

17 rLeuAsnThrLeuLeuLeu 23
|||||||:::|||||
51132 TCTAAATTCCTTTGATCTT 51150

seq_name: gb_htg22:AL451047

seq_documentation_block:
LOCUS AL451047 161545 bp DNA HTG 07-APR-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-1B20, *** SEQUENCING IN
PROGRESS ***, 28 unordered pieces.
ACCESSION AL451047 AC055798
VERSION AL451047.1 GI:11602617
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161545)
AUTHORS Burton,J.
TITLE Direct Submission
JOURNAL Submitted (06-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BALB20
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 146641 bases at least Q40
Consensus quality: 152102 bases at least Q30
Consensus quality: 155353 bases at least Q20
Insert size: 158845; sum-of-contigs
Insert size: 171731; agarose-fp
Quality coverage: 3.30x in Q20 bases; sum-of-contigs Quality
coverage: 3.29x in Q20 bases; agarose-fp

Draft Sequence Produced by Whitehead Institute/MIT Center for
Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA
http://www.seq.wi.mit.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4795: contig of 4795 bp in length
* 4796 4895: gap of 100 bp
* 4896 9059: contig of 4164 bp in length
* 9060 9159: gap of 100 bp
* 9160 12528: contig of 3369 bp in length
* 12529 12628: gap of 100 bp
* 12629 26952: contig of 14324 bp in length
* 26953 27052: gap of 100 bp
* 27053 31767: contig of 4715 bp in length
* 31768 31867: gap of 100 bp
* 31868 36136: contig of 4269 bp in length
* 36137 36236: gap of 100 bp
* 36237 39538: contig of 3302 bp in length
* 39539 39638: gap of 100 bp

* 39639 50902: contig of 11264 bp in length
* 50903 51002: gap of 100 bp
* 51003 57533: contig of 6531 bp in length
* 57534 57633: gap of 100 bp
* 57634 79585: contig of 21952 bp in length
* 79586 79685: gap of 100 bp
* 79686 85812: contig of 6127 bp in length
* 85813 85912: gap of 100 bp
* 85913 88090: contig of 2178 bp in length
* 88091 88190: gap of 100 bp
* 88191 99185: contig of 10995 bp in length
* 99186 99285: gap of 100 bp
* 99286 102918: contig of 3633 bp in length
* 102919 103018: gap of 100 bp
* 103019 105939: contig of 2921 bp in length
* 105940 106039: gap of 100 bp
* 106040 110442: contig of 4403 bp in length
* 110443 110542: gap of 100 bp
* 110543 114097: contig of 3555 bp in length
* 114098 114197: gap of 100 bp
* 114198 116474: contig of 2277 bp in length
* 116475 116574: gap of 100 bp
* 116575 123146: contig of 6572 bp in length
* 123147 123246: gap of 100 bp
* 123247 127668: contig of 4422 bp in length
* 127669 127768: gap of 100 bp
* 127769 130242: contig of 2474 bp in length
* 130243 130342: gap of 100 bp
* 130343 132879: contig of 2537 bp in length
* 132880 132979: gap of 100 bp
* 132980 136877: contig of 3898 bp in length
* 136878 136977: gap of 100 bp
* 136978 139814: contig of 2837 bp in length
* 139815 139914: gap of 100 bp
* 139915 147002: contig of 7088 bp in length
* 147003 147103: gap of 100 bp
* 147103 150773: contig of 3671 bp in length
* 150774 150873: gap of 100 bp
* 150874 158190: contig of 7317 bp in length
* 158191 158290: gap of 100 bp
* 158291 161545: contig of 3255 bp in length.
Location/Qualifiers
1. .161545 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-1B20"
/clone_lib="RPC1-11.1"
1. .4795 /note="assembly_fragment:00003.0"
4896. .9059 /note="assembly_fragment:00051"
9160. .12528 /note="assembly_fragment:00082"
12629. .26952 /note="assembly_fragment:00138"
27053. .31767 /note="assembly_fragment:00162"
31868. .36136 /note="assembly_fragment:00282"
36237. .39538 /note="assembly_fragment:00329"
39639. .50902 /note="assembly_fragment:00331"
51003. .57533 /note="assembly_fragment:00381"
57634. .79585 /note="assembly_fragment:00394"
79686. .85812 /note="assembly_fragment:00423"
85913. .88090 /note="assembly_fragment:00424"
88191. .99185

/note="assembly_fragment:00549"
99286. .102918 /note="assembly_fragment:00592"
103019. .105939 /note="assembly_fragment:00595"
106040. .110442 /note="assembly_fragment:00612"
110543. .114097 /note="assembly_fragment:00640"
114198. .116474 /note="assembly_fragment:00736"
116575. .123146 /note="assembly_fragment:00809"
123247. .127668 /note="assembly_fragment:00834"
127769. .130242 /note="assembly_fragment:00845"
130343. .132879 /note="assembly_fragment:00859"
132980. .136877 /note="assembly_fragment:00877"
136978. .139814 /note="assembly_fragment:00900"
139915. .147002 /note="assembly_fragment:00916"
147103. .150773 /note="assembly_fragment:00954"
150874. .158190 /note="assembly_fragment:01048"
158291. .161545 /note="assembly_fragment:01173"
BASE COUNT 51501 a 29748 c 28475 g 49116 t 2705 others
ORIGIN

alignment_scores:
Quality: 62.00 Length: 18
Ratio: 4.133 Gaps: 0
Percent Similarity: 83.333 Percent Identity: 66.667

alignment_block:
US-09-215-435-177_COPY_l_24 x AL451047/rev ..
Align seg 1/1 to reverse of: AL451047 from: 1 to: 161545

7 GlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuLe 23
|||||:||||| |||||||||||||||||||:|:|:|
9646 GGTTCCTTCCTTTATTTTATTTATTTCTGAGTTGAAATGGTGCCT 9597

23 ugly 24
|||||
9596 AGGT 9593

seq_name: gb_htg13:AC055798

seq_documentation_block:
LOCUS AC055798 167314 bp DNA HTG 15-JUN-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-1B20 map 1, WORKING DRAFT
SEQUENCE 42 unordered pieces.
ACCESSION AC055798
VERSION AC055798.3 GI:8567859
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 167314)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 1, clone RP11-1B20
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167314)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,

Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Rhmann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 15, 2000 this sequence version replaced gi:8113488.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7563

Center clone name: L.B.20

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 146335 bases at least Q40

Consensus quality: 156294 bases at least Q30

Consensus quality: 160582 bases at least Q20

Insert size: 176000; agarose-fp

Quality coverage: 3.2 in Q20 bases; agarose-fp

Quality coverage: 3.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 42 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1267: contig of 1267 bp in length

1268 1367: gap of 100 bp

1368 2781: contig of 1414 bp in length

2782 2881: gap of 100 bp

2882 3957: contig of 1076 bp in length

3958 4057: gap of 100 bp

4058 5450: contig of 1393 bp in length

5451 5550: gap of 100 bp

5551 6932: contig of 1382 bp in length

6933 7032: gap of 100 bp

7033 8658: contig of 1626 bp in length

8659 8758: gap of 100 bp

8759 10155: contig of 1397 bp in length

10156 10255: gap of 100 bp

10256 11921: contig of 1666 bp in length

11922 12021: gap of 100 bp

12022 13078: contig of 1057 bp in length

13079 13178: gap of 100 bp

13179 14458: contig of 1280 bp in length

* 14459 14558: gap of 100 bp

* 14559 15639: contig of 1081 bp in length

* 15640 15739: gap of 100 bp

* 15740 18149: contig of 2410 bp in length

* 18150 18249: gap of 100 bp

* 18250 21248: contig of 2999 bp in length

* 21249 21348: gap of 100 bp

* 21349 23819: contig of 2471 bp in length

* 23820 23919: gap of 100 bp

* 23920 27086: contig of 3167 bp in length

* 27087 27186: gap of 100 bp

* 27187 29461: contig of 2275 bp in length

* 29462 29561: gap of 100 bp

* 29562 32385: contig of 2824 bp in length

* 32386 32485: gap of 100 bp

* 32486 35779: contig of 3294 bp in length

* 35780 35879: gap of 100 bp

* 35880 39099: contig of 3220 bp in length

* 39100 39199: gap of 100 bp

* 39200 42626: contig of 3427 bp in length

* 42627 42726: gap of 100 bp

* 42727 46753: contig of 4027 bp in length

* 46754 46853: gap of 100 bp

* 46854 50733: contig of 3880 bp in length

* 50734 50833: gap of 100 bp

* 50834 54751: contig of 3918 bp in length

* 54752 54851: gap of 100 bp

* 54852 57785: contig of 2934 bp in length

* 57786 57885: gap of 100 bp

* 57886 60899: contig of 3014 bp in length

* 60900 60999: gap of 100 bp

* 61000 64635: contig of 3636 bp in length

* 64636 64735: gap of 100 bp

* 64736 68502: contig of 3767 bp in length

* 68503 68602: gap of 100 bp

* 68603 71649: contig of 3047 bp in length

* 71650 71749: gap of 100 bp

* 71750 76193: contig of 4444 bp in length

* 76194 76293: gap of 100 bp

* 76294 80893: contig of 4600 bp in length

* 80894 80993: gap of 100 bp

* 80994 87486: contig of 6493 bp in length

* 87487 87586: gap of 100 bp

* 87587 91852: contig of 4266 bp in length

* 91853 91952: gap of 100 bp

* 91953 97812: contig of 5860 bp in length

* 97813 97912: gap of 100 bp

* 97913 102283: contig of 4371 bp in length

* 102284 102383: gap of 100 bp

* 102384 108866: contig of 6483 bp in length

* 108867 108966: gap of 100 bp

* 108967 114480: contig of 5514 bp in length

* 114481 114580: gap of 100 bp

* 114581 121779: contig of 7199 bp in length

* 121780 121879: gap of 100 bp

* 121880 126681: contig of 4802 bp in length

* 126682 126781: gap of 100 bp

* 126782 133824: contig of 7043 bp in length

* 133825 133924: gap of 100 bp

* 133925 144194: contig of 10270 bp in length

* 144195 144294: gap of 100 bp

* 144295 155524: contig of 11230 bp in length

* 155525 155624: gap of 100 bp

* 15625 167314: contig of 11690 bp in length.

FEATURES

source

1. 167314
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-1B20"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature 1. 1267

```

/note="assembly_fragment"
1368. .2781
/note="assembly_fragment"
2882. .3957
/note="assembly_fragment"
4058. .5450
/note="assembly_fragment"
5551. .6932
/note="assembly_fragment"
7033. .8658
/note="assembly_fragment"
8759. .10155
/note="assembly_fragment"
10256. .11921
/note="assembly_fragment"
12022. .13078
/note="assembly_fragment"
13179. .14458
/note="assembly_fragment"
14559. .15639
/note="assembly_fragment"
15740. .18149
/note="assembly_fragment"
18250. .21248
/note="assembly_fragment"
21349. .23819
/note="assembly_fragment"
23920. .27086
/note="assembly_fragment"
27187. .29461
/note="assembly_fragment"
29562. .32385
/note="assembly_fragment"
32486. .35779
/note="assembly_fragment"
35880. .39099
/note="assembly_fragment"

alignment_scores:
Quality: 62.00 Length: 18
Ratio: 4.133 Gaps: 0
Percent Similarity: 83.333 Percent Identity: 66.667

alignment_block:
US-09-215-435-177_COPY_1_24 x AC055798/rev ..
Align seg 1/1 to reverse of: AC055798 from: 1 to: 167314

7 G1pHeLeuLeuArgPheHePheCysSerLeuAsnThrLeuLeuLe 23
|||||:|||| ||||||:|||||:|||||: ||
24205 GGTTTTTCTTATTTTTTATATTCGTAGTTGAAATGGTGCCT 24156

23 uGly 24
||||
24155 AGGT 24152

seg_name: gb_htg4:AC014616

seq_documentation_block:
LOCUS AC014616 120188 bp DNA HTG 16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC014616
VERSION AC014616.1 GI:6436719
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 120188)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission

```

Submitted (16-NOV-1999), Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10212336 by the submitter.
For further information on this sequence e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

Location/Qualifiers
1..120188
 organism="Drosophila melanogaster"
 /db_xref="taxon:7227"

BASE COUNT 39683 a 20457 c 20389 g 39659 t

ORIGIN

alignment_scores:
Quality: 60.00 Length: 20
Ratio: 3.750 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:
US-09-215-435-177_COPY1_1_24 x AC014616 ..

Align seg 1/1 to: AC014616 from: 1 to: 120188

2 LysSerAlaIysLeuGlyPheLeuArgPhePheillePheCysSerLe 18
|||||:||||||| ||||| |||:::||| ||
54820 AATAGTGTTGGTTTAGGCCATCTCTTCATTACATTTATTGTTCCT 54869

18 uAsnThrLeu 21
||||:||||
54870 AAATAGCCTA 54879

seq_name: gb_hgt2:AC010577

seq_documentation_block:
LOCUS AC010577 131684 bp DNA HTG 13-OCT-1999
DEFINITION Drosophila melanogaster chromosome 4 clone BACR22J20 (D1076)
RFBI-98 22.J.20 map 101F-102F strain y; cn bw sp, *** SEQUENCING IN
PROGRESS ***, 58 unordered pieces.
AC010577
AC010577.3 GI:6041726
HTG: HTGS_PHASE1.
fruit fly
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Drosophilidae; Drosophila

REFERENCE	1 (bases 1 to 131684)	Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G., Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleib, J.M., Park, S., Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E., Svirkas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rubin, G.M.
TITLE	Sequencing of Drosophila melanogaster	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 131684)	Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G., Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleib, J.M., Park, S., Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E., Svirkas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rubin, G.M.
TITLE	Direct Submission	
JOURNAL	Submitted (16-SEP-1999)	Drosophila Genome Center, Lawrence Berkeley

TITLE
Direct Submission
JOURNAL
Submitted (16-SEP-1999) Drosophila Genome Center, Lawrence Berkeley

COMMENT

Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Oct 15, 1999 this sequence version replaced gi:5922055.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bug@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 58 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
* 539: contig of 539 bp in length
* 619: gap of unknown length
* 540
* 1252: contig of 633 bp in length
* 1253
* 1332: gap of unknown length
* 1333
* 2031: contig of 699 bp in length
* 2032
* 2111: gap of unknown length
* 2112
* 2503: contig of 392 bp in length
* 2504
* 2583: gap of unknown length
* 2584
* 3686: contig of 1083 bp in length
* 3687
* 3747
* 4422: contig of 676 bp in length
* 4423
* 4502: gap of unknown length
* 4503
* 5556: contig of 1054 bp in length
* 5557
* 5636: gap of unknown length
* 5637
* 6287: contig of 651 bp in length
* 6288
* 6367: gap of unknown length
* 6368
* 7159: contig of 792 bp in length
* 7160
* 7239: gap of unknown length
* 7240
* 8535: contig of 1296 bp in length
* 8536
* 8615: gap of unknown length
* 8616
* 9729: contig of 1114 bp in length
* 9730
* 9809: gap of unknown length
* 9810
* 10558: contig of 749 bp in length
* 10559
* 10639
* 11610: contig of 972 bp in length
* 11611
* 11690: gap of unknown length
* 11691
* 12426: contig of 736 bp in length
* 12427
* 12506: gap of unknown length
* 12507
* 13606: contig of 1100 bp in length
* 13607
* 13686: gap of unknown length
* 13687
* 15581: contig of 1895 bp in length
* 15582
* 15661: gap of unknown length
* 15662
* 17324: contig of 1663 bp in length
* 17325
* 17404: gap of unknown length
* 17405
* 18936: contig of 1532 bp in length
* 18937
* 19016: gap of unknown length
* 19017
* 20108: contig of 1092 bp in length
* 20109
* 20188: gap of unknown length
* 20189
* 22142: contig of 1954 bp in length
* 22143
* 22222: gap of unknown length
* 22223
* 23815: contig of 1593 bp in length
* 23816
* 23895: gap of unknown length
* 23896
* 25279: contig of 1384 bp in length
* 25280
* 25359: gap of unknown length
* 25360
* 27692: contig of 2333 bp in length
* 27693
* 27772: gap of unknown length
* 29390: contig of 1618 bp in length
* 29391
* 29470: gap of unknown length
* 29471
* 31101: contig of 1631 bp in length
* 31102
* 31181: gap of unknown length
* 31182
* 33910: contig of 2729 bp in length
* 33911
* 33990: gap of unknown length
* 33991
* 35890: contig of 1900 bp in length
* 35891
* 35970: gap of unknown length
* 35971
* 38552: contig of 2582 bp in length
* 38553
* 38632: gap of unknown length
* 39807: contig of 1175 bp in length
* 39808
* 39887: gap of unknown length

* 39888
* 41904: contig of 2017 bp in length
* 41905
* 41984: gap of unknown length
* 45618: contig of 3634 bp in length
* 45698: gap of unknown length
* 45699
* 47674: contig of 1976 bp in length
* 47675
* 47734: gap of unknown length
* 47735
* 49814: contig of 1980 bp in length
* 49815
* 53262: contig of 3448 bp in length
* 53263
* 53437: gap of unknown length
* 53438
* 56427: contig of 3085 bp in length
* 56428
* 56507: gap of unknown length
* 56508
* 59066: contig of 2559 bp in length
* 59067
* 59146: gap of unknown length
* 59147
* 61633: contig of 2487 bp in length
* 61634
* 61713: gap of unknown length
* 61714
* 64993: contig of 3280 bp in length
* 64994
* 65074
* 67635: contig of 2582 bp in length
* 67636
* 67735: gap of unknown length
* 67736
* 70362: contig of 2627 bp in length
* 70363
* 70442: gap of unknown length
* 70443
* 74287: contig of 3845 bp in length
* 74288
* 74367: gap of unknown length
* 74368
* 80483: contig of 6116 bp in length
* 80484
* 80563: gap of unknown length
* 80564
* 86400: contig of 5837 bp in length
* 86401
* 86480: gap of unknown length
* 86481
* 95655: contig of 9175 bp in length
* 95656
* 95735: gap of unknown length
* 95736
* 104481: contig of 8746 bp in length
* 104482
* 114326: contig of 9765 bp in length
* 114327
* 114328: gap of unknown length
* 114329
* 124091: contig of 9685 bp in length
* 124092
* 124171: gap of unknown length
* 124172
* 124731: contig of 560 bp in length
* 124732
* 124811: gap of unknown length
* 124812
* 125551: contig of 740 bp in length
* 125552
* 125631: gap of unknown length
* 125632
* 126207: contig of 576 bp in length
* 126208
* 126287: gap of unknown length
* 126288
* 126904: contig of 617 bp in length
* 126905
* 126984: gap of unknown length
* 126985
* 127691: contig of 707 bp in length
* 127692
* 127771: gap of unknown length
* 127772
* 128367: contig of 596 bp in length
* 128368
* 128447: gap of unknown length
* 128448
* 129055: contig of 608 bp in length
* 129056
* 129135: gap of unknown length
* 129136
* 129672: contig of 537 bp in length
* 129673
* 129752: gap of unknown length
* 129753
* 130352: contig of 600 bp in length
* 130353
* 130432: gap of unknown length
* 130433
* 131056: contig of 624 bp in length
* 131057
* 131136: gap of unknown length
* 131137
* 131684: contig of 548 bp in length.

FEATURES

source

Location/Qualifiers

1. 131684
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="4"
/map="101F-102F"
/clone="BACR22J20 (D1076) RPCI-98 22.J.20"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial ECORI in
pBACE3.6")

BASE COUNT 41744 a 21575 c 22174 g 41614 t 4577 others
ORIGIN

alignment_scores:

Thu Nov 15 10:51:55 2001

Quality: 60.00 Length: 20
 Ratio: 3.750 Gaps: 0
 Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:

US-09-215-435-177_COPY_1_24 x AC010577 ..

Align seg 1/1 to: AC010577 from: 1 to: 131684

2 LysSerAlaLeuGlyPheLeuArgPhePheLeuPheCysSerLe 18

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

41209 AATGTGTTCTGTTAGCCATCTGCTTTCTTCTACATTTATTTCTCT 41258

18 uAsnThrLeu 21

|||||:|||||

41259 AATAGCCTA 41268

seq_name: gb_htg10:AC025616

seq_documentation_block:

LOCUS AC025616 137670 bp DNA HTG 20-APR-2000

DEFINITION Homo sapiens clone RP11-21J10, WORKING DRAFT SEQUENCE, 16 unordered pieces.

ACCESSION AC025616

VERSION AC025616.3 GI:7596801

KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 137670)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G.,

Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,

Collamore, A., Cooke, P., DeLlano, K., Dewar, K., Diaz, J. S.,

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hages, B., Heatford, A., Horton, L.,

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lechoczek, J.,

Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., Mlenga, V., Morrow, J.,

Melchior, J., Meneus, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

Murphy, T., Naylor, J., Oliver, J., Peterson, K., Pierle, N.,

O'Neil, D., Olivari, T. M., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

Pisani, C., Pollara, V., Raymond, C., Severy, P., Spencer, B.,

Roy, A., Santos, R., Schauer, S., Sever, P., Talamas, J.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Trifilio, J.,

Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trifilio, J.,

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,

Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 19, 2000 this sequence version replaced gi:7331406.

All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4030

Center clone name: 21_J_10

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 128269 bases at least Q40

Consensus quality: 133012 bases at least Q30

Consensus quality: 134872 bases at least Q20

Insert size: 136000; agarose-fp

Insert size: 136170; sum-of-contigs

Quality coverage: 4.2 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1388: contig of 1388 bp in length
 1389 1488: gap of 100 bp
 1489 3431: contig of 1943 bp in length
 3432 3531: gap of 100 bp
 3532 6098: contig of 2567 bp in length
 6099 6198: gap of 100 bp
 6199 9757: contig of 3559 bp in length
 9758 9857: gap of 100 bp
 9858 12836: contig of 2969 bp in length
 12837 12926: gap of 100 bp
 12927 16317: contig of 3391 bp in length
 16318 16417: gap of 100 bp
 16418 20454: contig of 4037 bp in length
 20455 20554: gap of 100 bp
 20555 24651: contig of 4097 bp in length
 24652 24751: gap of 100 bp
 24752 31538: contig of 6787 bp in length
 31539 31638: gap of 100 bp
 31639 39011: contig of 7373 bp in length
 39012 39111: gap of 100 bp
 39112 47730: contig of 8619 bp in length
 47731 47830: gap of 100 bp
 47831 61833: contig of 14003 bp in length
 61834 61933: gap of 100 bp
 61934 75049: contig of 13116 bp in length
 75050 75149: gap of 100 bp
 75150 91402: contig of 16253 bp in length
 91403 91502: gap of 100 bp
 91503 108120: contig of 16617 bp in length
 108120 108220: gap of 100 bp
 108220 137670: contig of 29451 bp in length.

FEATURES

Source

1. 137670
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-21J10"
 /clone_lib="RP11-11 Human Male BAC"

misc_feature

1. 1388

/note="assembly_fragment"

1489..3431

/note="assembly_fragment"

3532..6098

/note="assembly_fragment"

clone_end:SP6

vector_side:right"

6199..9757

/note="assembly_fragment"

9858..12826

/note="assembly_fragment"

12927..16317

/note="assembly_fragment"

16418..20454

/note="assembly_fragment"

20555..24651

/note="assembly_fragment"

Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Krawitz, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Krennison, S., Rulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Moberly, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. J., Pacleb, J. R., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spiet, E., Spradling, A. C., Turner, R., Watson, M., Strong, R. P., Sun, E., Svirskaas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wasserman, D. A., Weintraub, C. M., Weissenbach, J., Williams, S. M., Woodgate, T., Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhu, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M., and Venter, J. C.

20190606
20190606 (24047, 2185, 2135) (2006)
2 (bases 1 to 298469)
Adams,M.D., Celisnaker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7304344.
Location/Qualifiers
1..298469
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="4"
complement(1028..>1306)
/gene="CG17469"
/product="CT38623"
/db_xref="FlyBASE:FBan0017469"
/db_xref="FlyBASE:FBgn0040845"
complement(<1028..>1306)
/gene="CG17469"

```

/db_xref="FLYBASE:FBan0017469"
/db_xref="FlyBASE:FBgn0040845"
complement(1028..1192)
/gene="CG17469"
/note="CG17469 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0017469"
/db_xref="FlyBASE:FBgn0040845"
/protein_id="AAF59375.1"
/db_xref="GI:7304345"
/translation="MTESGIDLGCFDMEFDLINTLLNDNDNMDFLPNVTENNEEYELKS
SSRCIRHNEM"
join(3595..3608,3730..3887,3945..4042,4116..4365,
4417..4547,4611..4779,6348..6499,7372..7501,7555..7671)
/gene="CG17471"
/product="CT38625"
/db_xref="FLYBASE:FBan0017471"
/db_xref="FlyBASE:FBgn0039924"
<3595..>7671
/gene="CG17471"
/map="102E1-102E3"
/db_xref="FLYBASE:FBan0017471"
/db_xref="FlyBASE:FBgn0039924"
join(3750..3887,3945..4042,4116..4365,4417..4547,
4611..4779,6348..6499,7372..7501,7555..7671)
/gene="CG17471"
/note="CG17471 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0017471"
/db_xref="FlyBASE:FBgn0039924"
/protein_id="AAF59376.2"

```

```

/db_xref="GI:10728139"
/translation="MEKISSSQPRILKKHPRVHQVKLFRANEPILSVFMGIN
HTINELSHVINPMLPPDDPRVSKYIKVDHLEKNEMSPHFVKVCEPLVFNLRER
FGVDVYRESLTSPQIDSSKGAQYQSYDKFFLIKSLTSEIRMHAFILKQY
HPYVVRIGKTLPLPYLGMRYITVESVOYIVVRNWFSSHLTIHKFKLKGSTVDRE
ASEKLEKNLPTFKDNDIKQKVKLDIGKAKDKMLDTLSNDVLLTKLHIMDYSLIV
GVHOCVRAEEALQDNLITVGRSESECECSGERFATNTTGIKIRMLYLNCIA
EREYFTAIIDLVTQGVKKQAAKATVKYISVNDGISTCDPEQYAKRFLDFMDKA
IE"
complement(join(7685..7850,7901..7982,8045..8187,
9658..10140,10199..10336))
/gene="CG17467"
/product="CT38621"
/db_xref="FLYBASE:FBan0017467"
/db_xref="FLYBASE:FBgn0022341"
complement(<7685..>10336)
/gene="CG17467"
/map="102E1-102E3"
/db_xref="FLYBASE:FBan0017467"
/db_xref="FLYBASE:FBgn0022341"
complement(join(7848..7850,7901..7982,8045..8187,
9658..10140,10199..10297))
/gene="CG17467"
/note="CG17467 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0017467"
/db_xref="FLYBASE:FBgn0022341"
/protein_id="AAF59377.1"
/db_xref="GI:730437"
/translation="MKLDANIKSGLSNHVOYITLPDKKYIHNQKVOENSDEIVKSDS
KPEVKVATELSKFFKFFETYPSPAKNKRIFRMTPPREGVWVFPFPPSETNQOISST
LFNDILQKTKSTILNKREMEQKSDQKKKNPKLKCFTPPPEISHQFIRSDT
EESDSDEQNSDESEINPSNSVYNDKALLEAQSVARAKOLRAKRAKFKWNNLEIQ
BTEGRIDIVYQLSINSIESISAKAIRERFENMKKSETAMENPSKTGIRFV"
join(18584..18839,18906..19163,25083..25226,25280..26049,
26111..26402,27005..27124,27178..27740)
/gene="pho"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT39329"
/db_xref="FLYBASE:FBan0017743"
/db_xref="FLYBASE:FBgn0002521"
<18584..>27740
/gene="pho"
/note="CG17743"
/map="102E3-102E5"
/db_xref="FLYBASE:FBan0017743"
/db_xref="FLYBASE:FBgn0002521"
join(19069..19163,25083..25226,25280..26049,26111..26402,
27005..27124,27178..27325)
/gene="pho"
/note="pho gene product; Nucleotide sequence of the Celera
sequence differs from the published sequence for this
transcript."
/codon_start=1
/db_xref="FLYBASE:FBan0017743"
/db_xref="FLYBASE:FBgn0002521"
/protein_id="AAF59378.1"
/db_xref="GI:7304348"
/translation="MAYERFGIILQSEYDEIDGNTKVNQKNEKGNYDHLHRKNAF
DRIHSESKRGDNVINYNHENDKIRAADNIFSSKLKMPNNSYEMNCPKNIGYE
NOETSKVLTSNLSNNDINTESGVDKNSPFLTLTGILNSKGREREOKLVHKTM
EGFESVMTWASGSDSDYSGDQIVGASOLLKKEEFGIDGFTSQNKEYQKMSKFT
NAQLEAMPHPLNSVQIMDHLIKERGNLSQENNISERILSKTTLSPKEEPLLLPSSSIE
LVNETAAMPHNNHTLSNHTGTDGLHALPSSVPFRIGLHEGVNDCLSTISQSTHQD
NTDSTGCGEMNLSTVSYTNDKKIACPHKGCNKHFRDSAMRKHLLHTHGPFRVHCAE
CGKAFVSSKLRHQLVHTEKFPQCTFECCGRFSLDENLRTHTGDRFPVCPF
DACNKAFAOSTNLKSHLITAKARNVTSISGKSCSNAESNSQSDTSANYVKVLEQD
SVTENHVPFVYAD"
complement(join(33838..34674,34722..35332))
/gene="CG17461"
/product="CT38613"
/db_xref="FLYBASE:FBan0017461"

```

```

gene
/db_xref="FLYBASE:FBgn0039925"
complement(<33838..>35332)
/gene="CG17461"
/map="102E3-102E5"
/db_xref="FLYBASE:FBan0017461"
/db_xref="FLYBASE:FBgn0039925"
complement(join(33838..34674,34722..35255))
/gene="CG17461"
/note="CG17461 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0017461"
/db_xref="FLYBASE:FBgn0039925"
/protein_id="AAF59380.1"
/db_xref="GI:7304350"
/translation="MTNRYLALVLYLEIYNRIRDLNKNKNTVINIFLKLPGCI
GVSVTLTQTPVYVANCYDMLHFGNKNRVTAATLMNKNSSRSHRTITITLESPELN
SIGDAFGGICRGKLSLVLDLAGSBRQKTAQGDRLKEASQINLSALGNVSSLDV
GKAKHVPFRDLSKTLRLIQQYVNSLGSNGNTKLTWISCSPTDIHYDETISTLRYSRA
KNISNKPKNEDPKDARLQYQNEILYKMLQESQOIINKNDPNKIKKSPKLIQH
TNMSTKNVQIIDGRNCKASFKNNSLTLPNFPILQSEVQLOARSLDILKRSI
IGGERIDHDFELKEKHMARVAAQRHLSAIALSRVCKEDRDLLQGHVATITQEDIK
NDYIRKCKEKIKMLEMEVSDLNSEFQLDREYLDLIRNLGQVAFHQQLFKFTSSNPG
KFDNRW"
complement(join(50956..51220,52591..52741))

```

```

mRNA
alignment_scores:
Quality: 60.00 Length: 20
Ratio: 3.750 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000
alignment_block:
US-09-215-435-177_COPY_1_24 x AE003846 ...
Align seg 1/1 to: AE003846 from: 1 to: 298469
2 LysSerAlaIysLeuGlyPheLeuLeuLeuArgPhePheLeuPheCysSerLe 18
||| : : : : : ||| ||| ||| : : : : : ||| |||
109664 AAATGTGTCGTTAGCCATCGTCTTCTTCTACATTTATGTTCCCT 109713
18 uAsnThrLeu 21
||| : : : : : |||
109714 AAATAGCCTA 109723
seq_name: gb_htg7:AC021204
seq_documentation_block:
LOCUS AC021204 158357 bp DNA HTG 17-AUG-2000
DEFINITION Homo sapiens chromosome X clone RP11-143G3, WORKING DRAFT SEQUENCE,
15 unordered pieces.
ACCESSION AC021204
VERSION AC021204.6 GI:9838090
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 158357)
Waterston,R.H.
Direct Submission
Submitted (15-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:8954229.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----

```

Center project name: H_NH0143G03
----- Summary Statistics -----
Sequencing vector: M13; 78%
Chemistry: Dye-primer ET; 78% of reads
Sequencing vector: plasmid; 22%
Chemistry: Dye-terminator Big Dye; 22% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 145325 bases at least Q40
Consensus quality: 149562 bases at least Q30
Consensus quality: 152152 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 158286; sum-of-contigs
Quality coverage: 3.43 in Q20 bases; agarose-fp
Quality coverage: 3.70 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1539: contig of 1539 bp in length
* 1540 1639: gap of unknown length
* 1640 2801: contig of 1162 bp in length
* 2802 2901: gap of unknown length
* 2902 5208: contig of 2307 bp in length
* 5209 5308: gap of unknown length
* 5309 6887: contig of 1579 bp in length
* 6888 6987: gap of unknown length
* 6988 9583: contig of 2596 bp in length
* 9584 9683: gap of unknown length
* 9684 13291: contig of 3608 bp in length
* 13292 13391: gap of unknown length
* 13392 16412: contig of 3021 bp in length
* 16413 16512: gap of unknown length
* 16513 21055: contig of 4543 bp in length
* 21056 21155: gap of unknown length
* 21156 26232: contig of 5077 bp in length
* 26233 26332: gap of unknown length
* 26333 35392: contig of 9060 bp in length
* 35393 35492: gap of unknown length
* 35493 47474: contig of 11982 bp in length
* 47475 47574: gap of unknown length
* 47575 63901: contig of 16327 bp in length
* 63902 64001: gap of unknown length
* 64002 89634: contig of 25633 bp in length
* 89635 89734: gap of unknown length
* 89735 121751: contig of 32016 bp in length
* 121751 121850: gap of unknown length
* 121851 158357: contig of 36507 bp in length.

FEATURES
source
1. 158357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-143C3"
Location/Qualifiers
1. 158357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-143C3"

BASE COUNT 50000 a 30000 c 29276 g 47671 t 1410 others
ORIGIN

alignment_scores:
Quality: 59.50 Length: 22
Ratio: 3.719 Gaps: 1
Percent Similarity: 72.727 Percent Identity: 63.636

alignment_block:

US-09-215-435-177_COPY_1_24 x AC021204/rev ..

Align seg 1/1 to reverse of: AC021204 from: 1 to: 158357

8 PheLeuLeuArg.....PhePheIlePheCysSerLeuAs 19

9417 TTTTTCCTCCGAGGGGAGTCTCTTTTTCCTGCTGTTTTCCTCCTCAA 9368
|||||
|||||

19 nThrlLeuLeuGly 24
|:::|||||

9367 TGTTCCTGCTGTGGGG 9352
|:::|||||

seq_name: gb.htg8:AC022211

seq_documentation_block:

LOCUS AC022211 191151 bp DNA HTG 21-NOV-2000
DEFINITION Homo sapiens chromosome 17 clone RP11-649A18, WORKING DRAFT
SEQUENCE, 1 unordered pieces.

ACCESSION AC022211

VERSION AC022211.4 GI:11245626

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 191151)

AUTHORS Waterston,R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 191151)

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (26-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT On Nov 21, 2000 this sequence version replaced gi:7631098.

----- Genome Center -----
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----
Center project name: H_NH0649A18

----- Summary Statistics -----
Sequencing vector: M13; 77%
Sequencing vector: plasmid; 23%

Chemistry: Dye-primer ET; 72% of reads
Chemistry: Dye-terminator Big Dye; 28% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 190743 bases at least Q40

Consensus quality: 191007 bases at least Q30

Consensus quality: 191106 bases at least Q20

Insert size: 204000; agarose-fp

Insert size: 191151; sum-of-contigs

Quality coverage: 7.54 in Q20 bases; agarose-fp

Quality coverage: 8.05 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 191151: contig of 191151 bp in length.

FEATURES
Location/Qualifiers
1. 191151
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-649A18"
1. 191151
/note="assembly_name:Contig20
clone_end:SP6
vector_side:left"

source

misc_feature

BASE COUNT 48904 a 46522 c 47897 g 47828 t
ORIGIN

```
55023 AACTTTATTTAGGA 55008
|||||
seq_name: gb_pr9:HSJ612B15
seq_documentation_block:
LOCUS HSJ612B15 127820 bp DNA PRI 30-SEP-2000
Human DNA sequence from clone RP4-612B15 on chromosome 1p22.2-31.1.
Contains the (possibly pseudo) gene for a novel protein similar to
60S ribosomal protein L17 (RPL17), the gene for CGI-61, endophilin
B1 and KIAA0491, ESTs, STSs, GSSs and two CpG islands, complete
sequence.
ACCESSION AL049597 GI:9663352
VERSION AL049597.29 GI:9663352
KEYWORDS HTG: CGI-61; CpG island; endophilin B1; KIAA0491; RPL17.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 127820)
AUTHORS Wallis, J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Aug 2, 2000 this sequence version replaced gi:9211554.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP4-612B15 is from the library RPCI-4 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-612B15.
FEATURES
source
Location/Qualifiers
1..127820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.2-31.1"
/clone="RP4-612B15"
/clone_lib="RPCI-4"
repeat_region 1..476
/note="HERV9 repeat: matches 980..1445 of consensus"
repeat_region 477..1327
/note="LTR12 repeat: matches 1..1440 of consensus"
repeat_region 1328..4490
/note="HERV9 repeat: matches 1792..4954 of consensus"
repeat_region 4484..4919
/note="Harlequin repeat: matches 840..1271 of consensus"
repeat_region 4909..5191
/note="HERV17 repeat: matches 1255..1540 of consensus"
repeat_region 5166..5302
/note="Harlequin repeat: matches 1584..1720 of consensus"
repeat_region 5297..5937

alignment_scores:
Quality: 59.00 Length: 23
Ratio: 3.105 Gaps: 0
Percent Similarity: 82.609 Percent Identity: 52.174

alignment_block:
US-09-215-435-177_COPY_1_24 x AC022211
Align seg 1/1 to: AC022211 from: 1 to: 191151
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCys 17
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
105346 GTCAAGTGTGCAATCTTGCTCACTGCTCATCTTTTGTGTA 105395
17 rLeuAsnThrLeuLeuLeu 23
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
105396 CATAAATGGGGTTTGCTA 105414

seq_name: gb_pr1:AC005002
seq_documentation_block:
LOCUS AC005002 119841 bp DNA PRI 21-AUG-1999
DEFINITION Homo sapiens clone DJ378P09, complete sequence.
ACCESSION AC005002
VERSION AC005002.2 GI:5757541
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 119841)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 119841)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 119841)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Aug 21, 1999 this sequence version replaced gi:3212989.
FEATURES
source
Location/Qualifiers
1..119841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DJ378P09"
BASE COUNT 37083 a 23671 c 21587 g 37500 t
ORIGIN

alignment_scores:
Quality: 58.00 Length: 22
Ratio: 3.222 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 54.545

alignment_block:
US-09-215-435-177_COPY_1_24 x AC005002/rev ..
Align seg 1/1 to reverse of: AC005002 from: 1 to: 119841
3 SerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSerLeu 19
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
55073 TCGTATCCCTTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTGTTCTGTTCC 55024
19 nThrLeuLeuLeuGly 24
..
..
..
```


repeat_region /note="HBRV9 repeat: matches 7756. .8395 of consensus"
5349. .6034
repeat_region /note="MSTB repeat: matches 1352. .1437 of consensus"
6035. .6641
repeat_region /note="LNR12 repeat: matches 1. .671 of consensus"
6295. .6797
misc_feature /note="match: GSS: Em:AQ46353"
6793. .6969
repeat_region /note="MIR repeat: matches 18. .163 of consensus"
6970. .7262
repeat_region /note="AluSg1 repeat: matches 2. .293 of consensus"
7263. .7350
repeat_region /note="MIR repeat: matches 163. .258 of consensus"
7355. .7495
repeat_region /note="L2 repeat: matches 2559. .2705 of consensus"
complement(7658. .8075)
misc_feature /note="match: GSS: Em:AQ138017"
8194. .8491
repeat_region /note="L2 repeat: matches 1995. .2305 of consensus"
9635. .9924
repeat_region /note="AluSg/x repeat: matches 132. .291 of consensus"
9925. .10223
repeat_region /note="AluSc repeat: matches 1. .299 of consensus"
10252. .10943
repeat_region /note="346 copies 2 mer tt 58% conserved"
10502. .10783
misc_feature /note="random repeat. Single clone region. Restriction
digest (BAMHI and HINDIII) and pUC partners confirm
assembly."
10966. .11254
repeat_region /note="AluSg repeat: matches 1. .288 of consensus"
11637. .11929
repeat_region /note="AluSp repeat: matches 1. .295 of consensus"
12129. .12503
repeat_region /note="LNR18C repeat: matches 1. .387 of consensus"
12565. .13003
repeat_region /note="LNR37B repeat: matches 9. .468 of consensus"
13019. .13306
repeat_region /note="AluSx repeat: matches 1. .292 of consensus"
13727. .14008
repeat_region /note="AluSg repeat: matches 1. .284 of consensus"
14022. .14337
repeat_region /note="AluSg repeat: matches 1. .313 of consensus"
14400. .14449
repeat_region /note="25 copies 2 mer aa 76% conserved"
15656. .15771
repeat_region /note="L1MA10 repeat: matches 6212. .6322 of consensus"
15767. .15944
repeat_region /note="L1MC1 repeat: matches 6154. .6332 of consensus"
15966. .16245
repeat_region /note="L1MB8 repeat: matches 5781. .6078 of consensus"
16261. .16313
repeat_region /note="L1MA10 repeat: matches 6233. .6285 of consensus"
16346. .16538
repeat_region /note="L1MB8 repeat: matches 5595. .5791 of consensus"
16539. .16830
repeat_region /note="AluSx repeat: matches 1. .292 of consensus"
16831. .17074
repeat_region /note="L1MB8 repeat: matches 5364. .5595 of consensus"
17111. .17158
repeat_region /note="24 copies 2 mer tg 81% conserved"
17172. .17377
repeat_region /note="L1MB8 repeat: matches 5136. .5349 of consensus"
17474. .17778
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
18332. .18584
repeat_region /note="AluJb repeat: matches 14. .265 of consensus"
18665. .18769
repeat_region /note="MIR repeat: matches 103. .212 of consensus"
19024. .19170
repeat_region /note="L2 repeat: matches 2479. .2624 of consensus"
19771. .20967
repeat_region /note="L1MB2 repeat: matches 4967. .6168 of consensus"

repeat_region 20970. .21198
/note="MSTB repeat: matches 192. .426 of consensus"
21203. .21345
repeat_region /note="AluY repeat: matches 156. .298 of consensus"
21354. .21536
repeat_region /note="MSTB repeat: matches 1. .189 of consensus"
21537. .21966
repeat_region /note="L1MB2 repeat: matches 4538. .4969 of consensus"
21984. .22039
repeat_region /note="L1 repeat: matches 4545. .4599 of consensus"
22029. .22753
repeat_region /note="L1MB2 repeat: matches 3821. .4538 of consensus"
22746. .22957
misc_feature /note="match: GSS: Em:AQ279252"
complement(22749. .23381)
misc_feature /note="match: GSS: Em:A2040837"
22751. .22892
misc_feature /note="match: STS: Em:G06064"
22764. .23165
misc_feature /note="match: GSS: Em:AQ194597"
complement(22792. .23346)
gene /gene="dJ612B15.1"
CDS complement(22792. .23346)
/gene="dJ612B15.1"
/note="probably a pseudogene
match: proteins: Sw:P35266 Tr:O64803 Sw:P35267 Sw:O48557
Sw:P18621 Tr:Q9W3W8 Sw:P54033 Sw:O59423 Sw:O59794
Tr:Q9Z322 Sw:P05740 Sw:P46990 Sw:O28359 Sw:P37380
Sw:O26115 Sw:P24049 Tr:O14339
/codon_start=1
/evidence=not_experimental
/product="dJ612B15.1 (novel protein similar to 60S
ribosomal protein L17 (RPL17))"
/protein_id="CAC10392.1"
/db_xref="GI:10443242"
/translation="MVVYSIDPENPKSKSRGSLRVHFKDTRETAQAIKGMHIRKA
TYLKDVTLQKCIPTPRRYNGVGKCAQAKQWGTGRWPKRSAEFLLMFRNAEGT
EYKGLDVSLVYITQVNRKAPMLRRTYRAHGRINSYMSPPWHIEMILTEKEQIVPKP
EREVAQKKRISQKLLKKQKLMARK"
complement(22844. .23381)
misc_feature /note="match: GSS: Em:A2040921"
complement(23031. .23133)
misc_feature /gene="dJ612B15.1"
/note="match: GSS: Em:AQ808942"
23130. .23276
misc_feature /note="match: GSS: Em:AQ793153"
23381. .23598
repeat_region /note="L1MB2 repeat: matches 3608. .3834 of consensus"
23599. .23679
repeat_region /note="AluSc repeat: matches 228. .306 of consensus"
23683. .24713
repeat_region /note="L1MB2 repeat: matches 2580. .3637 of consensus"
24714. .24993
repeat_region /note="AluY repeat: matches 17. .302 of consensus"
24994. .25264
repeat_region /note="L1MB2 repeat: matches 2315. .2580 of consensus"

alignment_scores:

Quality: 58.00 Length: 24
Ratio: 3.412 Gaps: 0
Percent Similarity: 70.833 Percent Identity: 54.167

alignment_block:

US-09-215-435-177_COPY_1_24 x HSJ612B15/rev ..

Align seg 1/1 to reverse of: HSJ612B15 from: 1 to: 127820

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheCysSe 17
:::||||| |||||||||:::||||| ::|||

us-09-215-435-177_copy_1_24.rge

Thu Nov 15 10:51:55 2001

9378 TTAATACTACTTCAACTCAAAATCTTGCTCAGCATGTTCTCTATTGTTT 9329

17 rLeuAsnThrLeuLeuGly 24

|||||
|||
9328 CCTGAACACCAGATTATCTGGG 9307

OM of: US-09-215-435-177_COPY_1_24 to: N_Geneseq_0601:* out_format : pfs
Date: Nov 15, 2001 4:26 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnl_1/USPTO_spool/US09215435/runat.14112001.141105.2157/app_query.fasta.1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435@cgnl_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-215-435-177_COPY_1_24
Query length: 24
Database: N_Geneseq_0601:*
Database sequences: 730101
Database length: 313950809
Search time (sec): 1043.840000

score_list:	Sequence	Strd Orig	zScore	EScore	Len	Documentation
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	120.00	322.51	5.4e-10	267
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	120.00	322.38	5.5e-10	271
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	120.00	316.42	1.2e-09	526
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	120.00	316.20	1.2e-09	539
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	54.00	111.77	296.28	9813
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	54.00	69.51	6.5e+04	10821
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	52.00	124.72	56.25	1272
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	52.00	87.46	6.7e+03	80450
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	51.00	136.57	12.31	252
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	50.50	121.90	80.77	1109
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	50.50	115.35	187.14	2299
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	50.50	115.35	187.14	2299
	/cgnl_9/gcgdata/geneseq/NA1998.DAT:AA00455	+	50.00	113.97	223.44	2307
	/cgnl_9/gcgdata/geneseq/NA1998.DAT:AA00455	+	50.00	112.27	277.82	2787
	/cgnl_9/gcgdata/geneseq/NA1998.DAT:AA00455	+	50.00	96.24	2.2e+03	16593
	/cgnl_9/gcgdata/geneseq/NA2001.DAT:AA00455	+	49.50	123.31	67.44	702
	/cgnl_9/gcgdata/geneseq/NA2001.DAT:AA00455	+	49.00	127.26	40.62	389
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	49.00	120.28	99.44	846
	/cgnl_9/gcgdata/geneseq/NA1999.DAT:AA00455	+	49.00	119.29	112.97	945
	/cgnl_9/gcgdata/geneseq/NA1994.DAT:AA00455	+	49.00	108.76	435.77	3049
	/cgnl_9/gcgdata/geneseq/NA1993.DAT:AA00455	+	49.00	106.60	575.13	3879
	/cgnl_9/gcgdata/geneseq/NA1991.DAT:AA00455	+	49.00	105.88	630.83	4203
	/cgnl_9/gcgdata/geneseq/NA2001.DAT:AA00455	+	49.00	104.35	767.57	4983
	/cgnl_9/gcgdata/geneseq/NA1998.DAT:AA00455	+	49.00	101.49	1.1e+03	6850
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	49.00	97.22	1.9e+03	11014
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.50	112.78	260.20	1677
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	128.25	35.78	258
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	126.63	44.05	309
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	125.38	51.69	355
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	124.18	60.34	406
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	48.00	123.17	68.64	454
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	121.71	82.75	534
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	112.25	278.60	1531
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	111.86	292.69	1598
	/cgnl_9/gcgdata/geneseq/NA2000.DAT:AA00455	+	48.00	111.55	304.55	1654
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	48.00	111.55	304.55	1654
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	48.00	111.55	304.55	1654
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	48.00	111.55	304.55	1654
	/cgnl_9/gcgdata/geneseq/NA1997.DAT:AA00455	+	48.00	109.38	402.32	2106

/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AA00455 - 48.00 108.24 465.69
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AA00455 - 48.00 108.19 469.06
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA00455 - 48.00 104.91 714.16
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AA00455 - 48.00 103.98 804.19
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA00455 - 48.00 101.43 1.1e+03
seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA00455

seq_documentation_block:
ID AAC00455 standard; cDNA; 267 BP.
XX
AC AAC00455;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 453.
XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX

OS Homo sapiens.
XX

PN EP1033401-A2.
XX

PD 06-SEP-2000.
XX

PF 21-FEB-2000; 2000EP-0200610.
XX

PR 26-FEB-1999; 99US-0122487.
XX

PA (GEST) GENSET.
XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX

DR WPI; 2000-500381/45.
XX

DR P-PSDB; AAG00449.
XX

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT Claim 1; SEQ ID 453; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.

SQ Sequence 267 BP; 74 A; 57 C; 52 G; 84 T; 0 other;

alignment_scores:
Quality: 120.00 Length: 24
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_1_24 x AAC00455
Align seg 1/1 to: AAC00455 from: 1 to: 267
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheCys 17
|||||

US-09-215-435-177_COPY_1_24 x AAX51569

Align seg 1/1 to: AAX51569 from: 1 to: 271

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheCysSe 17
|||||
32 ATGAAGTCTGCCAAGCTGGGATTCTTCTTAAGATTCCTCATCTTGCTC 81

17 rLeuAsnThrLeuLeuGly 24
|||||
82 ATTGAATACCCCTGTATTGGGT 103

seq_name: /cgnl_9/gcgdata/geneseq/NA1999.DAT.AAX51569

seq_documentation_block:
ID AAX51569 standard; cDNA: 271 BP.
XX AC AAX51569;
XX DT 21-JUN-1999 (first entry)
XX DE Human secreted protein 5' EST SEQ ID NO:148.
XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX KW Homo sapiens.
XX OS
XX PN W09906549-A2.
XX PD 11-FEB-1999.
XX PF 31-JUL-1998; 98WO-IB01231.
XX PR 01-AUG-1997; 97US-0905279.
XX PA (GEST) GENSET.
XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX WP1: 1999-153779/13.
XX P-PSDB; AAY12791.
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
XX Claim 1; Page 253-254; 522pp; English.
XX
XX AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12681 to
CC AAY12913, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoietic regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

Seq Sequence 271 BP; 75 A; 57 C; 55 G; 84 T; 0 other;

alignment_scores:
Quality: 120.00 Length: 24
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_1_24 x AAX97612 ..

Align seg 1/1 to: AAX97612 from: 1 to: 526

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSe 17
|||||
22 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTCTTCATCTCTGCTC 71
17 rleuAsnThrLeuLeuLeuGly 24
|||||
72 ATTGAATACCCCTGTATTGGGT 93

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX97738

seq_documentation_block:
ID AAX97738 standard; DNA; 539 BP.

XX
AC AAX97738;
XX
DT 13-SEP-1999 (first entry)
DE
DE Extended human secreted protein coding sequence, SEQ ID NO. 303.
XX
XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.
XX
OS Homo sapiens.
XX
XX WO9931236-A2.
XX
XX 24-JUN-1999.
XX
XX 17-DEC-1998; 98WO-IB02122.
XX
XX 10-AUG-1998; 98US-0096116.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
XX
XX (GEST) GENSET.
XX
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
XX WPI; 1999-385906/32.
DR P-PSDB; AAY36054.
XX
XX New isolated human secreted proteins
XX
XX Claim 1; Page 376; 516pp; English.
XX
XX This sequence represents an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX
XX Sequence 539 BP; 165 A; 103 C; 97 G; 162 T; 12 other;

alignment_scores:
Quality: 120.00 Length: 24
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_1_24 x AAX97738 ..

Align seg 1/1 to: AAX97738 from: 1 to: 539

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSe 17
|||||
32 ATGAAGTCTGCCAAGCTGGGATTTCTTCTAAGATTCTTCATCTCTGCTC 81
17 rleuAsnThrLeuLeuLeuGly 24
|||||
82 ATTGAATACCCCTGTATTGGGT 103

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX13059

seq_documentation_block:
ID AAX13059 standard; DNA; 9813 BP.

XX
AC AAX13059;
XX
DT 19-MAR-1999 (first entry)
DE
DE Enterococcus faecalis genome contig SEQ ID NO:122.
XX
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
XX
OS Enterococcus faecalis.
XX
XX WO9850555-A2.
XX
XX 12-NOV-1998.
XX
XX 04-MAY-1998; 98WO-US08985.
XX
XX 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Barash SC, Dillon PJ, Kunsch CA;
PI WPI; 1999-045171/04.
XX
XX New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
XX
XX Claim 1; Page 734-739; 2084pp; English.
XX
XX A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AAX12938 to AAX13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
XX
XX Sequence 9813 BP; 3233 A; 1698 C; 2161 G; 2719 T; 2 other;

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:

Quality: 52.00

Length: 19

Ratio: 3.467

Gaps: 0

Percent Similarity: 78.947

Percent Identity: 52.632

us-09-215-435-177_copy_1_24.rng

Thu Nov 15 10:51:55 2001

```

alignment_block:
US-09-215-435-177_COPY_1_24 x AAC34882  ..
Align seg 1/1 to: AAC34882 from: 1 to: 1272

5 LysLeuGlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLe 21
||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:  |||
135 AAAATGGGTTTATAACATCGATTTTATGTTTTCATGCTCTTCACACAG 184

21 uleuLeu 23
||||:
185 TTGGTT 191

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAF22295
seq_documentation_block:
ID AAF22295 standard; DNA: 80450 BP.
XX
AC AAF22295;
XX
DT 20-MAR-2001 (first entry)
XX
DE BAC containing repeats from centromeres 1-4 #18.
XX
KW Centromere; michrosome; vector; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200055325-A2.
XX
PR 21-SEP-2000.
XX
PS (UYCH-) UNIV CHICAGO.
XX
PI Preuss D, Copenhaver G, Keith K;
XX
DR WPI; 2000-587529/55.
XX
PR 18-MAR-1999; 99US-0125219.
XX
PR 01-APR-1999; 99US-0127409.
XX
PR 18-MAY-1999; 99US-0134770.
XX
PR 13-SEP-1999; 99US-0153584.
XX
PR 17-SEP-1999; 99US-0154603.
XX
PS Claim 102; Page 668-686; 1449pp; English.
XX
CC The present invention relates to a recombinant DNA construct of a plant
XX (Arabidopsis thaliana) centromere. The constructs are useful for
XX producing stably inherited michrosomes which can serve as vectors for
XX the construction of transgenic plant and animal cells expressing
XX selected proteins such as hormones, enzymes, interleukins, clotting
XX factors, cytokines, antibodies, and growth factors.
XX
SQ Sequence 80450 BP; 25014 A; 14662 C; 14629 G; 26145 T; 0 other;

alignment_scores:
Quality: 52.00 Length: 22
Ratio: 2.889 Gaps: 1
Percent Similarity: 81.818 Percent Identity: 50.000

alignment_block:
US-09-215-435-177_COPY_1_24 x AAF22295/rev ..
Align seg 1/1 to reverse of: AAF22295 from: 1 to: 80450

5 LysLeuGlyPheLeuLeuArg.....PhePheIlePheCysSerLeuAs 19
||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:  |||
31464 AAGCTTGCCTTCATGGTTAAGAACTGCTTCATGGAAGCTTCTGCTTTGAG 31415

19 nThrLeuLeuGly 24
||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:  |||
31414 GAGGCTTATCTCTGGC 31399

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA221845
seq_documentation_block:
ID AA221845 standard; cDNA: 252 BP.
XX
AC AA221845;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 25920.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PS (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID 25920; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
SQ Sequence 252 BP; 72 A; 49 C; 64 G; 66 T; 1 other;

alignment_scores:
Quality: 51.00 Length: 16
Ratio: 3.643 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 56.250

alignment_block:
US-09-215-435-177_COPY_1_24 x AA221845 ..
Align seg 1/1 to: AA221845 from: 1 to: 252

7 GlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeu 22
||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:  |||

```


161 GGCCTTCCTGAGATTTTTCATGATGACAGAAATGCAATTCCTG 208

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF72760

seq_documentation_block:

ID AAF72760 standard; cDNA; 1109 BP.

XX AAF72760;

AC AAF72760;

XX 24-APR-2001 (first entry)

DT Human prostate cancer antigen coding sequence #20.

DE Immunosuppressive; nontropic; neuroprotective; antiviral; vulnery;

XX anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant;

KW immune disorder; cardiovascular disorder; neurological disease;

KW infection; cancer; cytostatic; antileukemic; antirheumatic;

KW antiasthmatic; anticonvulsant; vasotropic; vulnery; human;

KW secreted protein; prostate cancer antigen; ss.

XX Homo sapiens.

OS WO200107476-A1.

XX 01-FEB-2001.

XX 20-JUL-2000; 2000WO-US19666.

XX 21-JUL-1999; 99US-0144972.

PR 13-AUG-1999; 99US-0148681.

PR 17-AUG-1999; 99US-0149173.

PR 06-OCT-1999; 99US-0158004.

PR 05-APR-2000; 2000US-0194689.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Rosen CA, Birse C;

XX WPI: 2001-138554/14.

DR P-PSDB; AAB80292.

XX New nucleic acid molecule encoding human secreted prostate cancer

PT antigens, useful for the diagnosis and treatment of disorders such as

PT cancer, leukemia and autoimmune disease .

XX Claim 1; Page 369; 433pp; English.

PS The present invention relates to human secreted prostate cancer antigen

CC coding sequences (AAF72741-AAF72789) and proteins (AAB80273-AAB80321).

CC The coding sequences and proteins of the present invention are useful for

CC preventing, treating or ameliorating a medical condition; and for the

CC diagnosis and treatment of diseases and disorders. Diseases and disorders

CC that can be diagnosed and treated include (auto)immune diseases (e.g.

CC graft versus host disease and rheumatoid arthritis), inflammatory and

CC allergic disorders (e.g. asthma), hyperproliferative disorders (e.g.

CC cancers and leukemias), cardiovascular disorders (e.g. heart attacks and

CC arrhythmias), cerebrovascular disorders (e.g. stroke), arterial occlusive

CC disorders (e.g. arteriosclerosis), angiogenesis related disorders (e.g.

CC retinopathy and keloid scars), ocular disorders (e.g. glaucoma),

CC neurological disorders (e.g. Alzheimer's, Parkinson's disease, epilepsy

CC and Creutzfeld-Jakob disease) and infections caused by bacteria, fungi,

CC viruses or parasites. They may also be useful for wound healing,

CC epithelial cell proliferation, supporting cell culture, tissue

CC regeneration, birth control and as a food additive or preservative.

CC The coding sequences can be used to generate fusion proteins by linking

CC the coding sequences to the human immunoglobulin G Fc portion coding

CC sequence (AAF72732) for increasing the stability of the fusion protein

CC as compared to the human protein only.

XX Sequence 1109 BP; 346 A; 211 C; 198 G; 354 T; 0 other;

alignment_scores:

Quality: 50.50 Length: 23

Ratio: 2.658 Gaps: 1

Percent Similarity: 82.609 Percent Identity: 52.174

alignment_block:

US-09-215-435-177_COPY_1_24 x AAF72760/rev ..

Align seg 1/1 to reverse of: AAF72760 from: 1 to: 1109

1 MetLysSerAlaLysLeu...GlyPheLeuLeuArgPhePheIlePheCy 16

||||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||

434 ATGAGTTCAGATAAAATGTTCAAAATATTTGCTAGATTTTAAATCGGTG 385

16 sSerLeuAsnThrLeuLeu 22

||||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||

384 CTCATGAGAACCCCTCATA 366

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF16129

seq_documentation_block:

ID AAF16129 standard; cDNA; 2299 BP.

XX AAF16129;

AC AAF16129;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:564.

DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;

KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease; ss.

XX Homo sapiens.

OS WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI: 2000-587513/55.

DR P-PSDB; AAB56926.

XX Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of

PT disorders such as prostate cancer .

XX Claim 1; Page 1022-1023; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytostatic,

CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,

CC nephrotropic, antiinfective, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

us-09-215-435-177_copy_1_24.rng

Thu Nov 15 10:51:55 2001

```

CC  invention.
XX  Sequence 2299 BP; 666 A; 437 C; 417 G; 774 T; 5 other;
SQ  Sequence 2299 BP; 666 A; 437 C; 417 G; 774 T; 5 other;

alignment_scores:
  Quality: 50.50      Length: 23
  Ratio: 2.658       Gaps: 1
Percent Similarity: 82.609 Percent Identity: 52.174

alignment_block:
US-09-215-435-177_COPY_1_24 x AAF16129/rev ..
Align seg 1/1 to reverse of: AAF16129 from: 1 to: 2299

1 MetLysSerAlaLysLeu...GlyPheLeuLeuArgPhePheLeuPheCy 16
1652 ATGAGTTCAGATAAAATGCTCAAAATATTTGCTTAGATTTTAAATCGGTG 1603

16 sSerLeuAsnThrLeuLeu 22
|||||:|||||:|||||:
1602 CTCATGAGAACCCCTCATA 1584

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF72785
seq_documentation_block:
ID AAF72785 standard; cDNA; 2299 BP.
XX AC AAF72785;
XX AC AAF72785;
XX DT 24-APR-2001 (first entry)
XX DE Human prostate cancer antigen coding sequence #45.
XX XX
KW Immunosuppressive; nootropic; neuroprotective; antiviral; vulnary;
KW anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac;
KW immune disorder; cardiovascular disorder; neurological disease;
KW infection; cancer; cytostatic; antiarthritic; antirheumatic;
KW antiasthmatic; anticonvulsant; vasotropic; vulnary; human;
KW secreted protein; prostate cancer antigen; ss.
XX OS Homo sapiens.
XX WO200107476-A1.
XX PN
XX PD 01-FEB-2001.
XX PF 20-JUL-2000; 2000WO-US19666.
XX PR 21-JUL-1999; 99US-0144972.
XX PR 13-AUG-1999; 99US-0148681.
XX PR 17-AUG-1999; 99US-0149173.
XX PR 06-OCT-1999; 99US-0158004.
XX PR 05-APR-2000; 2000US-0194689.
XX XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX FA
XX PI Rosen CA, Birse C;
XX XX
XX WPI; 2001-138554/14.
XX DR P-PSDB; AAB80317.
XX XX
XX New nucleic acid molecule encoding human secreted prostate cancer
XX PT antigens, useful for the diagnosis and treatment of disorders such as
XX PT cancer, leukemia and autoimmune disease.
XX PS
XX Claim 1; Page 382; 433pp; English.
XX CC
XX The present invention relates to human secreted prostate cancer antigen
XX CC coding sequences (AAF72741-AAF72789) and proteins (AAB80273-AAB80321). for
XX CC The coding sequences and proteins of the present invention are useful for
XX CC preventing, treating or ameliorating a medical condition; and for the

```

```

CC diagnosis and treatment of diseases and disorders. Diseases and disorders
CC that can be diagnosed and treated include (auto)immune diseases (e.g.
CC graft versus host disease and rheumatoid arthritis), inflammatory and
CC allergic disorders (e.g. asthma), hyperproliferative disorders (e.g.
CC cancers and leukemias), cardiovascular disorders (e.g. heart attacks and
CC arrhythmias), cerebrovascular disorders (e.g. stroke), arterial occlusive
CC disorders (e.g. arteriosclerosis), angiogenesis related disorders (e.g.
CC retinopathy and keloid scars), ocular disorders (e.g. glaucoma),
CC neurological disorders (e.g. Alzheimer's, Parkinson's disease, epilepsy
CC and Creutzfeld-Jakob disease) and infections caused by bacteria, fungi,
CC viruses or parasites. They may also be useful for wound healing,
CC epithelial cell proliferation, supporting cell culture, tissue
CC regeneration, birth control and as a food additive or preservative.
CC The coding sequences can be used to generate fusion proteins by linking
CC the coding sequences to the human immunoglobulin G Fc portion coding
CC sequence (AAF72732) for increasing the stability of the fusion protein
CC as compared to the human protein only.
XX
XX Sequence 2299 BP; 666 A; 437 C; 417 G; 774 T; 5 other;
SQ

```

```

alignment_scores:
  Quality: 50.50      Length: 23
  Ratio: 2.658       Gaps: 1
Percent Similarity: 82.609 Percent Identity: 52.174

alignment_block:
US-09-215-435-177_COPY_1_24 x AAF72785/rev ..
Align seg 1/1 to reverse of: AAF72785 from: 1 to: 2299

1 MetLysSerAlaLysLeu...GlyPheLeuLeuArgPhePheLeuPheCy 16
1652 ATGAGTTCAGATAAAATGCTCAAAATATTTGCTTAGATTTTAAATCGGTG 1603

16 sSerLeuAsnThrLeuLeu 22
|||||:|||||:|||||:
1602 CTCATGAGAACCCCTCATA 1584

```

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV43009

```

```

seq_documentation_block:
ID AAV43009 standard; DNA; 2307 BP.
XX AC AAV43009;
XX AC AAV43009;
XX DT 09-NOV-1998 (first entry)
XX DE Streptococcus pneumoniae polypeptide coding region.
XX KW Polypeptide; ORF; open reading frame; infection; bacterial;
XX KW streptococcal; bacteremia; diagnosis; prophylaxis; ds.
XX OS Streptococcus pneumoniae.
XX PH Key Location/Qualifiers
XX CDS 1878..2306
XX FT /*tag= a
XX FT /note= "polypeptide"
XX XX
XX WO9823631-A1.
XX PN
XX PD 04-JUN-1998.
XX XX
XX 24-NOV-1997; 97WO-US21976.
XX PR 27-NOV-1996; 96US-0031879.
XX XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
XX PI Reid RH, zarfos PN;

```

```
XX WPI; 1998-322654/28.
DR P-PSDB; AAW62734.
XX Streptococcus pneumoniae polynucleotides - useful for developing
PT products for diagnosis, prevention and treatment of infections e.g.
PT pneumonia, bacteremia, meningitis or endocarditis
XX
XX Claim 1; Page 127; 181pp; English.
PS The sequence is that of a Streptococcal polypeptide coding region.
CC The polypeptide can potentially be used for the diagnosis and
CC prevention of bacterial infections, especially SP infection.
CC It may be used for the treatment of diseases such as otitis media,
CC conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural
CC empyema, endocarditis or infection of the cerebrospinal fluid.
XX
SQ Sequence 2307 BP; 742 A; 362 C; 509 G; 694 T; 0 other;

alignment_scores:
  Quality: 50.00 Length: 17
  Ratio: 3.571 Gaps: 0
Percent Similarity: 82.353 Percent Identity: 47.059

alignment_block:
US-09-215-435-177_COPY_1_24 x AAV43009 ..
Align seg 1/1 to: AAV43009 from: 1 to: 2307

7 GlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuLe 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1610 GGATTTTGGATGGAAGTGCATGCTGTTTTCGCTTTTAAACAGAAATTTAAT 1659

23 u 23
1660 G 1660

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ96363
seq_documentation_block:
ID AAZ96363 standard; DNA; 2787 BP.
XX
XX AAZ96363;
AC
XX
XX 10-APR-2000 (first entry)
DT
XX
XX S. pneumoniae derived DNA from ORF #191.
DE
XX Treatment; prevention; disease; diagnosis; gene therapy; screening;
KW bacterial; antimicrobial; antibiotic; pathogenesis; infection; ss.
XX
XX Streptococcus pneumoniae.
OS
XX
XX WO9806734-A1.
PN
XX
XX 19-FEB-1998.
PD
XX
XX 15-AUG-1997; 97WO-US14436.
PF
XX
XX 16-AUG-1996; 96US-0024022.
PR
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA
XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
PI Stodola RK;
PI
XX
XX WPI; 1998-159452/14.
DR
XX P-PSDB; AAY86028, AAY86029, AAY86030.
DR
XX Streptococcus pneumoniae proteins and related DNA - useful for
PT screening compounds for antibacterial activity
XX
```

```
PS Claim 4; Page 222-223; 640pp; English.
XX
XX This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see AAZ96173-296494) and their encoded proteins (see
CC AAY8792-Y86183). The DNA, vectors and host cells described in the
CC method of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of
CC disease, or diagnosis of disease related to expression or activity of
CC such a polypeptide. They can also be used to screen for compounds which
CC interact with and inhibit or activate such a polypeptide. The
CC polypeptides (or DNA encoding them, via gene therapy) are also useful
CC for inducing an immunological response in a mammal. The antagonists are
CC useful to inhibit such bacterial polypeptides. The polypeptides are
CC particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
CC infection, dysfunction and disease.
XX
SQ Sequence 2787 BP; 883 A; 449 C; 607 G; 848 T; 0 other;

alignment_scores:
  Quality: 50.00 Length: 17
  Ratio: 3.571 Gaps: 0
Percent Similarity: 82.353 Percent Identity: 47.059

alignment_block:
US-09-215-435-177_COPY_1_24 x AAZ96363 ..
Align seg 1/1 to: AAZ96363 from: 1 to: 2787

7 GlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuLe 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2090 GGATTTTGGATGGAAGTGCATGCTGTTTTCGCTTTTAAACAGAAATTTAAT 2139

23 u 23
2140 G 2140

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAV52185
seq_documentation_block:
ID AAV52185 standard; DNA; 16593 BP.
XX
XX AAV52185;
AC
XX
XX 23-OCT-1998 (first entry)
DT
XX
XX Streptococcus pneumoniae genome fragment SEQ ID NO:52.
DE
XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
XX
XX Streptococcus pneumoniae.
OS
XX
XX WO9818931-A2.
PN
XX
XX 07-MAY-1998.
PD
XX
XX 30-OCT-1997; 97WO-US19588.
PF
XX
XX 31-OCT-1996; 96US-0029960.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI Kunsch CA, Rosen CA;
PI
XX
XX WPI; 1998-272225/24.
DR
XX Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
```


OM of: US-09-215-435-177_COPY_1_24 to: Issued_Patents_NA.* out_format : pfs
Date: Nov 15, 2001 3:58 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=framet_p2n.model -DEV=xlp
-O=/cgnl_1/uspro.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB=Issued_Patents_NA -QPMW=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -XGAPEXT=0.000
-GAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPEXT=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pcp
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@cgnl_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-215-435-177_COPY_1_24
Query length: 24
Database: Issued_Patents_NA.*
Database sequences: 351203
Database length: 113238999
Search time (sec): 228.030000

score_list:	Strd Orig	zScore	EScore Len	Documentation
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-943-731-2 +	50.00	91.03	2.0e+03	38682 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-298-367B-2 -	49.00	104.48	360.55	6850 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-991-426-3 +	48.00	115.02	93.33	1654 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-143-470-3 +	48.00	115.02	93.33	1654 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-396-479B-7 -	48.00	111.50	146.64	2406 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-818-823-7 -	48.00	111.50	146.64	2406 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-219-932-5 +	48.00	107.10	257.75	3841 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-623-906A-3 -	47.00	128.93	150.68	279 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-949-246-3 +	47.00	112.38	130.90	1622 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-07-635-561A-7 +	46.50	132.63	9.75	162 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-248-335-57 -	46.00	114.49	99.88	960 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-329-796-2 -	46.00	112.54	128.22	1181 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-415-593-44 -	46.00	111.27	150.93	1352 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-123-343A-6 -	46.00	101.31	541.36	3900 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-257-073-6 -	45.50	113.02	120.59	966 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-602-359A-27 -	45.00	112.34	131.62	894 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-899-324-32 +	45.00	102.19	483.88	2632 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-329-892B-32 +	45.00	102.19	483.88	2632 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-605-106-7 -	45.00	99.92	647.20	3350 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-605-106-7 -	45.00	95.32	1.2e+03	5467 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-394-189B-1 -	45.00	92.16	1.8e+03	7653 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-258-287B-1 -	45.00	92.16	1.8e+03	7653 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-368-704C-1 -	45.00	92.16	1.8e+03	7653 !
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-05701-18 -	45.00	92.16	1.8e+03	7653 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-285-440-23 -	44.00	120.27	47.63	285 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-630-349-23 -	44.00	120.27	47.63	285 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-385-982-229 +	44.00	113.11	119.22	610 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-408-095-26 +	44.00	108.99	202.34	946 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-100-391-5 -	44.00	107.00	261.17	1169 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-100-391-11 -	44.00	105.98	297.40	1302 !
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US91-08177-18 -	44.00	102.54	462.84	1879 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-100-391-9 -	44.00	99.33	698.35	2643 !
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-248-466B-2 -	44.00	98.91	736.75	2763 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-972-927-2 +	44.00	86.88	3.4e+03	9936 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-966-958-1 -	44.00	86.14	3.8e+03	10754 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-215-817-1 -	44.00	86.14	3.8e+03	10754 !
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-342-353-1 -	44.00	86.14	3.8e+03	10754 !
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-721-925A-4 +	43.00	116.24	79.83	324 !
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-714-918-47 +	43.00	112.49	129.20	483 !

/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-265-315-47 + 43.00 112.49 129.20 483
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-265-315-47 + 43.00 112.49 129.20 483
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-266-417-47 + 43.00 112.49 129.20 483
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-958-629C-1 - 43.00 107.67 239.54 806
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US94-06661-1 - 43.00 100.85 574.89 16
seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-943-731-2

seq documentation block:
; Sequence 2, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KOKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-943-731-2

alignment_scores:
Quality: 50.00 Length: 20
Ratio: 3.33 Gaps: 0
Percent Similarity: 75.000 Percent Identity: 60.000
alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-943-731-2

```

; FILING DATE: 16-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/851,160
; FILING DATE: 05-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/643,798
; FILING DATE: 07-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 86...1276
; US-08-991-426-3

alignment_scores:
    Quality: 48.00      Length: 24
    Ratio: 2.824        Gaps: 0
    Percent Similarity: 70.833      Percent Identity: 45.833

alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-991-426-3 ..
Align seg 1/1 to: US-08-991-426-3 from: 1 to: 1654

1 MetLysSerAlaLysLeuGlyPheLeuArgPhePheLeuArgPhePheCysSe 17
||| :||||:||||:||||:||||:||||:||||:||||:||||:||||:
86 ATGGCTCCGATATCTCTGCTGGCTGCTCCGCTTGGCCACCTTCTGCCA 135

17 rLeuAsnThrLeuLeuGly 24
||||:||||:||||:||||:
136 TCTGACTGCTCTGCTGGCTGGA 157

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-143-470-3

seq_documentation_block:
; Sequence 3, Application US/09143470
; Patent No. 6043086
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NEUROTACTIN AND USES THEREFOR
; FILE REFERENCE: 09404/049001
; CURRENT APPLICATION NUMBER: US/09/143,470
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)...(1276)
; US-09-143-470-3

alignment_scores:
    Quality: 48.00      Length: 24
    Ratio: 2.824        Gaps: 0
    Percent Similarity: 70.833      Percent Identity: 45.833

```

```

; FILING DATE: 16-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/851,160
; FILING DATE: 05-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/643,798
; FILING DATE: 07-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 86...1276
; US-08-991-426-3

alignment_scores:
    Quality: 48.00      Length: 24
    Ratio: 2.824        Gaps: 0
    Percent Similarity: 70.833      Percent Identity: 45.833

alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-991-426-3 ..
Align seg 1/1 to: US-08-991-426-3 from: 1 to: 1654

1 MetLysSerAlaLysLeuGlyPheLeuArgPhePheLeuArgPhePheCysSe 17
||| :||||:||||:||||:||||:||||:||||:||||:||||:||||:
86 ATGGCTCCGATATCTCTGCTGGCTGCTCCGCTTGGCCACCTTCTGCCA 135

17 rLeuAsnThrLeuLeuGly 24
||||:||||:||||:||||:
136 TCTGACTGCTCTGCTGGCTGGA 157

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-143-470-3

seq_documentation_block:
; Sequence 3, Application US/09143470
; Patent No. 6043086
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NEUROTACTIN AND USES THEREFOR
; FILE REFERENCE: 09404/049001
; CURRENT APPLICATION NUMBER: US/09/143,470
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)...(1276)
; US-09-143-470-3

alignment_scores:
    Quality: 48.00      Length: 24
    Ratio: 2.824        Gaps: 0
    Percent Similarity: 70.833      Percent Identity: 45.833

```

```

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-143-470-3
..
Align seg 1/1 to: US-09-143-470-3 from: 1 to: 1654

1 MetLysSerAlaLysLeuGlyPheLeuArgPheIlePheCysSe 17
|||||:|||||:|||||:|||||:|||||
86 ATGGCTCGGATATCTCTCGTGGTGCTCGCTGGCCACCTTCTGCCA 135

17 rleuAsnThrLeuLeuGly 24
|||||:|||||:|||||
136 TCTGACTGCTCGTGGCTGGA 157

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-396-479B-7

seq_documentation_block:
; Sequence 7, Application US/08396479B
; Patent No. 5612455
; GENERAL INFORMATION:
; APPLICANT: HOEY, Timothy
; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/396,479B
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59450-1/RAO
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 210 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2406 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 211..2337
; US-08-396-479B-7

alignment_scores:
Quality: 48.00 Length: 15
Ratio: 3.429 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 60.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-396-479B-7/rev ..
Align seg 1/1 to reverse of: US-08-396-479B-7 from: 1 to: 2406

8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeu 22
|||||:|||||:|||||:|||||:|||||
2401 TTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 2357

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-09-219-932-5

seq_documentation_block:
; Sequence 5, Application US/09219932B
; Patent No. 6153199
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BUBLOT, Michel
; APPLICANT: RIVIERE, Michel

```

```

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-818-823-7

seq_documentation_block:
; Sequence 7, Application US/08818823
; Patent No. 5708158
; GENERAL INFORMATION:
; APPLICANT: HOEY, Timothy
; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,823
; FILING DATE: 14-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/396,479
; FILING DATE: 02-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59450-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 210 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2406 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 211..2337
; US-08-818-823-7

alignment_scores:
Quality: 48.00 Length: 15
Ratio: 3.429 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 60.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-818-823-7/rev ..
Align seg 1/1 to reverse of: US-08-818-823-7 from: 1 to: 2406

8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeu 22
|||||:|||||:|||||:|||||:|||||
2401 TTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTT 2357

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-219-932-5

seq_documentation_block:
; Sequence 5, Application US/09219932B
; Patent No. 6153199
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BUBLOT, Michel
; APPLICANT: RIVIERE, Michel

```

TITLE OF INVENTION: AVIAN RECOMBINANT VACCINE USING, A
 TITLE OF INVENTION: INFECTIOUS LARYNGOTRACHEITIS VIRUS
 FILE REFERENCE: 454313-2210
 CURRENT APPLICATION NUMBER: US/09/219,932B
 CURRENT FILING DATE: 1998-12-23
 EARLIER APPLICATION NUMBER: PCT/FR97/01138
 EARLIER FILING DATE: 1997-06-25
 EARLIER APPLICATION NUMBER: 96/08243
 EARLIER FILING DATE: 1996-06-27
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 3841
 TYPE: DNA
 ORGANISM: Infectious Laryngotracheitis Virus
 09-09-219-932-5

```

alignment_scores:
  Quality: 48.00
  Ratio: 4.364
  Percent Similarity: 78.571
  Length: 14
  Gaps: 0
  Percent Identity: 71.429

alignment_block:
  US-09-215-435-177_COPY_1_24 x US-09-219-932-5 ..
  Align seq 1/1 to: US-09-219-932-5 from: 1 to: 3841

6 LeuGlyPheLeuLeuArgPhePheLeuPheCysSerLeuAsn 19
      ||||| ||||| ||||| ||||| ||||| |||||
2115 TTGGATTTCAACTCCGACGGGGCATTTTGTGCAGGATAAC 2156
      ||||| ||||| ||||| ||||| ||||| |||||

```

6 LeuGlyPheLeuLeuArgPhePheIlePheCysSerLeuAsn 19
|||||
2115 TTGGGATTCAACTCCGACCGGGCCATTTTCTGCAGCATAAAC 2156

US-08-623-90

```
seq_documentation_block:
  . Sequence 3 Application US/08623906A
```

sequence of application and ;
; Patent No. 5874217
; GENERAL INFORMATION:
;

APPLICANT: Stevenson, Tamara
APPLICANT: Dvorak, Jan

APPLICANT: Halverson, Joy
TITLE OF INVENTION: Microsatellite Sequences for Canine

```

;
; TITLE OF INVENTION: Genotyping
; NUMBER OF SEQUENCES: 60
;

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, suite 3400
CITY: San Francisco

```

;
; STATE: CA
; COUNTRY: US
;

```

```
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; *****
```

```

;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
;
; OPERATING SYSTEM: PC-DOS/MS DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; ----- PROTECTION DATA: -----
```

```

:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/623,906A
:
: FILING DATE:

```

CLASSIFICATION: 435
FILING DATE:
ATTORNEY / AGENT INFORMATION:

NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36.677

REGISTRATION NUMBER: A-62282/BIR
REFERENCE/DOCKET NUMBER: A-62282/BIR
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 279 base pairs
TYPE: nucleic acid

```

STRANDEDNESS: double
TOPOLOGY: linear

```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1622 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-949-246-3

alignment_scores:
  Quality: 47.00      Length: 12
  Ratio: 4.273       Gaps: 0
Percent Similarity: 91.667      Percent Identity: 75.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-949-246-3 ..
Align seg 1/1 to: US-08-949-246-3 from: 1 to: 1622

12 PhePheIlePheCysSerLeuAsnThrLeuLeu 23
|||||
1456 TTTTTCATCTTTTGTCAATTAAACCTTGTAGTG 1491

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-635-561A-7

seq_documentation_block:
; Sequence 7, Application US/07635561A
; Patent No. 5244805
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; TITLE OF INVENTION: Improved Baculovirus Expression Vectors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/635,561A
; FILING DATE: 19910117
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/353,847
; FILING DATE: 17-MAY-1989
; PRIOR APPLICATION DATA: WO 90/02814
; FILING DATE: 17-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 4-90A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; TELEX: 823189
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 92..160
; OTHER INFORMATION: /note= "mcs #2 - Multicloning site
```

```
; OTHER INFORMATION: #2"
US-07-635-561A-7

alignment_scores:
  Quality: 46.50      Length: 27
  Ratio: 2.583       Gaps: 1
Percent Similarity: 66.667      Percent Identity: 40.741

alignment_block:
US-09-215-435-177_COPY_1_24 x US-07-635-561A-7 ..
Align seg 1/1 to: US-07-635-561A-7 from: 1 to: 162

1 MetLysSerAlaLysLeuGly.....PheLeuLeuArgPh 12
|||||
2 ATGTCCCGGCCACGCTTGGCGTTATTGAATAAGAAATTTAAAAATCAATC 51

12 ePheIlePheCysSerLeuAsnThrLeuLeu 22
|||||
52 ATTTGTATACTGTAATTAATACATACCTGTTT 82

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-248-335-57

seq_documentation_block:
; Sequence 57, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 57
; LENGTH: 960
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-57

alignment_scores:
  Quality: 46.00      Length: 17
  Ratio: 3.067       Gaps: 0
Percent Similarity: 88.235      Percent Identity: 41.176

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-248-335-57/rev ..
Align seg 1/1 to reverse of: US-09-248-335-57 from: 1 to: 960

8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeu 24
|||||
913 TTTTTTTTTTTTTTTTTTTTTTTGTAGTAGATGATCGATTACGATTGG 864

24 y 24
863 A 863

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-329-796-2

seq_documentation_block:
; Sequence 2, Application US/09329796
; Patent No. 6291174
; GENERAL INFORMATION:
; APPLICANT: Ning Li
; APPLICANT: Changxin WU
; APPLICANT: Yaofeng ZHAO
; TITLE OF INVENTION: DNA MARKERS FOR PIG LITTER SIZE
```

74

alignment_scores:
Quality: 46.00 Length: 12
Ratio: 5.111 Gaps: 0
Percent Similarity: 75.000 Percent Identity: 75.000
alignment_block:
US-09-215-435-177_COPY_1_24 x US-08-123-343A-6/rev ..
Align seg 1/1 to reverse of: US-08-123-343A-6 from: 1 to: 3900
8 PheLeuLeuArgPhePheHeillePheCysSerLeuAsn 19
|||||
2354 TTTTGTGCGTTTTTTCGCTTTGTTTCAACAAT 2319
seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-257-073-6
seq_documentation_block:
; Sequence 6, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-257-073-6
alignment_scores:
Quality: 45.50 Length: 15
Ratio: 3.792 Gaps: 1
Percent Similarity: 80.000 Percent Identity: 73.333
alignment_block:

US-09-215-435-177_COPY_1_24 x US-08-257-073-6/rev ..
Align seg 1/1 to reverse of: US-08-257-073-6 from: 1 to: 966
8 PheLeuLeuArgPhePheHeillePheCysSerLeuAsnThrLeuLeu 22
|||||
559 TTCTGCTTCGATTCCTT...TTCTGTTTCTTCATCTCTCTGCTT 518

us-09-215-435-177_copy_1_24.rni

Thu Nov 15 10:51:55 2001

OM of: US-09-215-435-177_COPY_1_24 to: Pending_Patents_NA_Main:* out_format : pfs
Date: Nov 15, 2001 3:54 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2173/app_query.fasta_1.1519
-DB=Pending_Patents_NA_Main -QPMT=fastap -SUFFIX=rnmpm
-GAPEXT=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-XGAPEXT=0.000 -XGAPEXT=4.500 -XGAPEXT=0.050 -XGAPEXT=10.000
-XGAPEXT=0.500 -XGAPEXT=6.000 -XGAPEXT=7.000 -XGAPEXT=10.000
-XGAPEXT=0.500 -XGAPEXT=6.000 -XGAPEXT=7.000 -XGAPEXT=10.000
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_12482 -NCPU=6 -ICPU=3 -LONGLOG
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY_1_24
Query length: 24
Database: Pending_Patents_NA_Main:*
Database sequences: 17159718
Database length: 173266264
Search time (sec): 11750.060000

score_list:	Strd Orig	zScore	Escore Len	Documentation
/cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-148 +	120.00	301.60	1.9e-07	271
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-3799 +	120.00	236.13	3.8e-07	499
/cgnl_7/ptodata/1/pna/US0619_COMB.seq:US-09-197-873-5799 +	120.00	296.13	3.8e-07	49
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-76 +	120.00	295.66	4.0e-07	526
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-303 +	120.00	295.44	4.1e-07	539
/cgnl_7/ptodata/1/pna/US0606_COMB.seq:US-06-069-957-136 +	120.00	295.44	4.1e-07	539
/cgnl_7/ptodata/1/pna/US0617_COMB.seq:US-06-171-625-457 +	60.00	109.79	9.0e+03	264
/cgnl_7/ptodata/1/pna/US0617_COMB.seq:US-06-173-464-18981 +	60.00	109.79	9.0e+03	2
/cgnl_7/ptodata/1/pna/US0619_COMB.seq:US-06-191-637-23127 +	60.00	109.79	9.0e+03	2
/cgnl_7/ptodata/1/pna/US0619_COMB.seq:US-06-191-681-18240 +	60.00	109.79	9.0e+03	2
/cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-528-237A-2004 -	60.00	92.74	8.0e+04	1
/cgnl_7/ptodata/1/pna/US096A_COMB.seq:US-09-619-643-31295 +	56.00	145.68	90.25	1
/cgnl_7/ptodata/1/pna/US096C_COMB.seq:US-09-654-617-350711 +	56.00	145.68	90.25	1
/cgnl_7/ptodata/1/pna/US096E_COMB.seq:US-09-684-016-350711 +	56.00	145.68	90.25	1
/cgnl_7/ptodata/1/pna/US0601A_COMB.seq:US-06-146-905-1329 +	56.00	145.68	90.25	15
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-209-830-49802 +	56.00	134.11	397.97	15
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-229-511-170 -	56.00	97.85	4.2e+04	327
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-229-511-171 -	56.00	97.84	4.2e+04	327
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-233-939-74 -	56.00	97.84	4.2e+04	327
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-226-176-2299 +	56.00	87.33	1.6e+05	10
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-233-468-2299 +	56.00	87.33	1.6e+05	10
/cgnl_7/ptodata/1/pna/US0603A_COMB.seq:US-06-313-371-2299 +	56.00	87.33	1.6e+05	10
/cgnl_7/ptodata/1/pna/US0604A_COMB.seq:US-06-248-498-61 -	56.00	80.66	3.8e+05	2225
/cgnl_7/ptodata/1/pna/US0604A_COMB.seq:US-06-248-542-62 -	56.00	80.66	3.8e+05	2225
/cgnl_7/ptodata/1/pna/US0604A_COMB.seq:US-06-245-227-9 -	56.00	80.54	3.8e+05	22553
/cgnl_7/ptodata/1/pna/US091_COMB.seq:US-09-112-430-837 -	55.00	138.46	227.74	267
/cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-478-821-837 -	55.00	138.46	227.74	267
/cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-229-188417 -	55.00	138.46	227.74	267
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-866-555-4358 +	55.00	134.42	382.42	419
/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-371-805-2173 -	55.00	131.75	538.26	564
/cgnl_7/ptodata/1/pna/US096B_COMB.seq:US-09-620-608-1790 -	55.00	123.26	1.6e+03	14
/cgnl_7/ptodata/1/pna/US0601A_COMB.seq:US-06-144-883-1790 -	55.00	123.26	1.6e+03	14
/cgnl_7/ptodata/1/pna/US0603A_COMB.seq:US-06-038-697-568 -	55.00	115.70	4.2e+03	337
/cgnl_7/ptodata/1/pna/US0604A_COMB.seq:US-06-046-714-561 -	55.00	115.70	4.2e+03	337
/cgnl_7/ptodata/1/pna/US096B_COMB.seq:US-09-620-392-17378 -	55.00	95.66	5.5e+04	3
/cgnl_7/ptodata/1/pna/US097A_COMB.seq:US-09-702-134-25324 +	55.00	95.66	5.5e+04	3
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-213-168-40 -	55.00	95.32	5.8e+04	3276
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-229-512-80 +	55.00	95.32	5.8e+04	3276
/cgnl_7/ptodata/1/pna/US0602A_COMB.seq:US-06-229-512-125 +	55.00	95.32	5.8e+04	327

/cgnl_7/ptodata/1/pna/US6023_COMB.seq:US-60-234-446-588 - 55.00 95.32 5.8e+04
/cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-528-237A-516 - 55.00 92.73 8.0e+04
/cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-534-859-775 - 55.00 87.92 1.5e+05
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-803-736-775 - 55.00 87.92 1.5e+05
/cgnl_7/ptodata/1/pna/US6025_COMB.seq:US-60-257-537-284 + 55.00 79.56 4.3e+05
seq_name: /cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-148

seq_documentation_block:
; Sequence 148, Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; TITLE OF INVENTION: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/905,279
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011A
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq PFIFCSLNTLLG/GV
US-08-905-279-148

alignment_scores:
Quality: 120.00 Length: 24
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-177_COPY_1_24 x US-08-905-279-148
Align seg 1/1 to: US-08-905-279-148 from: 1 to: 271
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheCys 17
|||||
32 ATGAAGTCTGCAAGCTGGGATTTCTTCTAAGATCTTCATCTTCTGTC 81
17 rLeuAsnThrLeuLeuGly 24
|||||

Thu Nov 15 10:51:56 2001

us-09-215-435-177_copy_1_24.rnmpm

```
82 ATTGAATACCGTGTATTGGGT 103
seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-5799
seq_documentation_block:
; Sequence 5799, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81 US2 REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 5799
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..318
; NAME/KEY: sig_peptide
; LOCATION: 22..93
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.59999990463257
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
US-09-834-366-5799

alignment_scores:
Quality: 120.00 Length: 24
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-197-873-5799 ..
Align seg 1/1 to: US-60-197-873-5799 from: 1 to: 499

1 MetLysSerAlaLysLeuGlyPheLeuArgPheHeilePheCysSe 17
|||||
22 ATGAAGTCTGCCAAGCTGGGATTCTCTAAGATTCTTCATCTCTGCTC 71

17 rLeuAsnThrLeuLeuGly 24
|||||
72 ATTGAATACCGTGTATTGGGT 93

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-76
seq_documentation_block:
; Sequence 76, Application US/09215435A
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bougueleret Lydie
; TITLE OF INVENTION: Extended CDNAS
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-2-9
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-4-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-8-10
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: Patent.pm
; SEQ ID NO 76
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..318
; NAME/KEY: sig_peptide
; LOCATION: 22..93
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 497..502
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 516..526
US-09-215-435-76
```

```
seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-5799
seq_documentation_block:
; Sequence 5799, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81 US1 PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 5799
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```

alignment_scores:
  Quality: 120.00      Length: 24
  Ratio: 5.000         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-177_COPY_1_24 x US-09-215-435-76  ..
  Align seg 1/1 to: US-09-215-435-76 from: 1 to: 526

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17
22 ATGAAGTCTGCCAAGCTGGATTCTTCTAAGATTCTTCATCTTCTGCTC 71
17 rLeuAsnThrLeuLeuGly 24
17 rLeuAsnThrLeuLeuGly 24
72 ATTGAATACCTGTTATTGGGT 93

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-303

seq_documentation_block:
; Sequence 303, Application US/09215435A
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bouqueleret Lydie
; TITLE OF INVENTION: Extended cDNAs
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-2-9
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-4-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-8-10
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: Patent.pm
; SEQ ID NO 303
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..328
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.5999990463257
; OTHER INFORMATION: seq PFIFCSLNTLLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 508..513
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 528..539
; US-09-215-435-303

alignment_scores:
  Quality: 120.00      Length: 24
  Ratio: 5.000         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-177_COPY_1_24 x US-09-215-435-303  ..
  Align seg 1/1 to: US-09-215-435-303 from: 1 to: 539

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSe 17

```

```

17 rLeuAsnThrLeuLeuGly 24
82 ATTGAATACCTGTTATTGGGT 103

seq_name: /cgnl_7/ptodata/1/pna/US6006_COMB.seq:US-60-069-957-136

seq_documentation_block:
; Sequence 136, Application US/60069957
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: EXTENDED cDNAs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 381
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/069,957
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.019PR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 136:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..328
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq PFIFCSLNTLLLG/GV
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 508..513
; FEATURE:
; NAME/KEY: poly_a
; LOCATION: 528..539
; IDENTIFICATION METHOD: blastn
; US-60-069-957-136

alignment_scores:
  Quality: 120.00      Length: 24
  Ratio: 5.000         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

us-09-215-435-177_copy_1_24.rnpgm

Thu Nov 15 10:51:56 2001

```

alignment_block:
US-09-215-435-177_COPY_1_24 x US-60-069-957-136  ..
Align seg 1/1 to: US-60-069-957-136 from: 1 to: 539

1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSe 17
|||||
32 ATGAAGTCGCGAAGTCGGATTCTTCTAAGATTCTTCATCTTGCTC 81
|||||
17 rLeuAsnThrLeuLeuGly 24
|||||
82 ATTGAATACCTGTTATGGGT 103

seq_name: /cgnl_7/ptodata/1/pna/US6017_COMB.seq:US-60-171-625-457

seq_documentation_block:
; Sequence 457, Application US/60171625
; GENERAL INFORMATION:
; APPLICANT: Skupski, Marian
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS THAT SHOW
; TITLE OF INVENTION: HIGH HOMOLOGY TO KNOWN HUMAN DISEASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000179
; CURRENT APPLICATION NUMBER: US/60/171,625
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 26499
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-171-625-457

alignment_scores:
Quality: 60.00 Length: 20
Ratio: 3.750 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-60-171-625-457  ..
Align seg 1/1 to: US-60-171-625-457 from: 1 to: 26499

2 LysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSerLe 18
|||||
23740 AAATGTCGTTAGGCCATCTGCTTTCTTCTACATTATTGTTCTCT 23789
|||||
18 uAsnThrLeu 21
|||||
23790 AAATAGCCTA 23799

seq_name: /cgnl_7/ptodata/1/pna/US6017_COMB.seq:US-60-173-464-18981

seq_documentation_block:
; Sequence 18981, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18981
; LENGTH: 26499
; TYPE: DNA
; ORGANISM: Drosophila
US-60-173-464-18981

alignment_scores:
Quality: 60.00 Length: 20
Ratio: 3.750 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-60-191-637-23127  ..
Align seg 1/1 to: US-60-191-637-23127 from: 1 to: 26500

2 LysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSerLe 18
|||||
23741 AAATGTCGTTAGGCCATCTGCTTTCTTCTACATTATTGTTCTCT 23790
|||||
18 uAsnThrLeu 21
|||||
23791 AAATAGCCTA 23800

seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-191-681-18240

seq_documentation_block:
; Sequence 18240, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF.
; FILE REFERENCE: CL000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18240
; LENGTH: 26500

```



```
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-18240

alignment_scores:
  Quality: 60.00      Length: 20
  Ratio: 3.750        Gaps: 0
  Percent Similarity: 80.000  Percent Identity: 55.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-60-191-681-18240 from: 1 to: 26500
Align seg 1/1 to: US-60-191-681-18240 from: 1 to: 26500
2 LysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSerLe 18
||| :|||:||||| ||||| |||||:||||| ||
23741 AAATGCTTCGTTAGGCCATCTGCTTCTTCTACATTTATTGTTTCCT 23790
18 uAsnThrLeu 21
|||:||||
23791 AAATAGCCTA 23800

seq_name: /cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-528-237A-2004
seq_documentation_block:
; Sequence 2004, Application US/09528237A
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
; TITLE OF INVENTION: Sequences and Uses Thereof
; FILE REFERENCE: CL000284
; CURRENT APPLICATION NUMBER: US/09/528,237A
; CURRENT FILING DATE: 2000-03-17
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 2926
; SEQ ID NO 2004
; LENGTH: 177295
; TYPE: DNA
; ORGANISM: Drosophila
US-09-528-237A-2004

alignment_scores:
  Quality: 60.00      Length: 20
  Ratio: 3.750        Gaps: 0
  Percent Similarity: 80.000  Percent Identity: 55.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-528-237A-2004/rev ..
Align seg 1/1 to reverse of: US-09-528-237A-2004 from: 1 to: 177295
2 LysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSerLe 18
||| :|||:||||| ||||| |||||:||||| ||
66807 AAATGCTTCGTTAGGCCATCTGCTTCTTCTACATTTATTGTTTCCT 66758
18 uAsnThrLeu 21
|||:||||
66757 AAATAGCCTA 66748

seq_name: /cgnl_7/ptodata/1/pna/US096A_COMB.seq:US-09-619-643-31295
seq_documentation_block:
; Sequence 31295, Application US/09619643
; GENERAL INFORMATION:
; APPLICANT: Fisher, Dane K.
; APPLICANT: Lalgudi, Raghunath V.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51230)B
; CURRENT APPLICATION NUMBER: US/09/619,643

; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-18240

; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 32236
; SEQ ID NO 31295
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3181-011-P1-K2-H4
US-09-619-643-31295

alignment_scores:
  Quality: 56.00      Length: 16
  Ratio: 4.308        Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 75.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-619-643-31295 ..
Align seg 1/1 to: US-09-619-643-31295 from: 1 to: 158
8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuLeu 23
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 TTCCTGTAAGTTTTTTTATTCTGTCCTGTTGTTGCTTTTCTT 147

seq_name: /cgnl_7/ptodata/1/pna/US096C_COMB.seq:US-09-654-617-350711
seq_documentation_block:
; Sequence 350711, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 350711
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-350711

alignment_scores:
  Quality: 56.00      Length: 16
  Ratio: 4.308        Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 75.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-654-617-350711 ..
Align seg 1/1 to: US-09-654-617-350711 from: 1 to: 158
8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuLeu 23
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 TTCCTGTAAGTTTTTTTATTCTGTCCTGTTGTTGCTTTTCTT 147

seq_name: /cgnl_7/ptodata/1/pna/US096E_COMB.seq:US-09-684-016-350711
seq_documentation_block:
; Sequence 350711, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 350711
; LENGTH: 158
```

us-09-215-435-177_copy_1_24.rnmpm

Thu Nov 15 10:51:56 2001

```

; TYPE: DNA
; ORGANISM: Zea mays
US-09-684-016-350711

alignment_scores:
    Quality: 56.00      Length: 16
    Ratio: 4.308        Gaps: 0
    Percent Similarity: 81.250      Percent Identity: 75.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-684-016-350711 ..
Align seg 1/1 to: US-09-684-016-350711 from: 1 to: 158
      8 PheLeuLeuArgPheIlePheCysSerLeuAsnThrLeuLeuLeu 23
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      100 TCTGTGTTAGTATTTTATTTCTGTTCTGTTCTGTTCTTCTT 147

seq_name: /cgnl_7/ptodata/1/pna/US6014_COMB.seq:US-60-146-905-1329

seq_documentation_block:
; Sequence 1329, Application US/60146905
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Deng, Mollan
; APPLICANT: Fisher, Dane K.
; APPLICANT: Miller, Phillip W.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(51232)A
; CURRENT APPLICATION NUMBER: US/60/146,905
; CURRENT FILING DATE: 1999-08-02
; NUMBER OF SEQ ID NOS: 2850
; SEQ ID NO 1329
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3181-011-P1-K2-H4
US-60-146-905-1329

alignment_scores:
    Quality: 56.00      Length: 16
    Ratio: 4.308        Gaps: 0
    Percent Similarity: 81.250      Percent Identity: 75.000

alignment_block:
US-09-215-435-177_COPY_1_24 x US-60-146-905-1329 ..
Align seg 1/1 to: US-60-146-905-1329 from: 1 to: 158
      8 PheLeuLeuArgPheIlePheCysSerLeuAsnThrLeuLeuLeu 23
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      100 TCTGTGTTAGTATTTTATTTCTGTTCTGTTCTGTTCTTCTT 147
```

OM of: US-09-215-435-177_COPY_1_24 to: Pending_Patents_NA_New:* out_format : pfs
Date: Nov 15, 2001 4:08 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n_model -DEV=xlp
-O=/cgnl_1/USPTO_spool/US09215435/runat_14112001_141105_2194/app_query.fasta_1.1519
-DB=Pending_Patents_NA_New -QFWT=fastap -SUFFIX=rnnp
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -SPART=1
-MATRIX=blosom62 -TRANS=human40.cgi -LIST=45 -DOCALIGN=200
-THR_SEQ=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09215435 -CGNL_1_69 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY_1_24
Query length: 24
Database: Pending_Patents_NA_New:*
Database sequences: 127876
Database length: 122380375
Search time (sec): 564.420000

score.list:

Sequence	Strd Orig	zScore	EScore	Len	! Documentation
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-969-708-319 +	53.00	128.73	5.90		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-8488 +	51.00	128.46	6.11		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-9659 +	51.00	126.75	7.61		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-11087 +	51.00	126.21	8.15		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-22094 +	51.00	125.45	8.99		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-22403 +	51.00	124.94	9.59		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-18975 -	51.00	124.88	9.67		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-9295 +	50.00	112.73	45.95		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-9530 +	50.00	112.73	45.95		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-214 +	50.00	104.43	133.23		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-9149 +	49.00	125.81	8.58		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-10026 +	49.00	121.22	15.46		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-5052 +	48.00	127.75	6.69		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-4662 -	47.00	114.28	37.66		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-111 +	47.00	111.88	51.26		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1444 +	47.00	111.88	51.26		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1030 -	46.00	115.10	33.91		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1758 -	46.00	115.10	33.91		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-760 -	46.00	105.82	111.49		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-973-278-875 +	46.00	104.78	127.31		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-694 -	46.00	101.75	187.82		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-937-974-25 -	46.00	97.46	325.65		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-795-668-667 -	45.00	109.93	65.78		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-795-668-668 -	45.00	109.93	65.78		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-5520 -	45.00	109.44	70.04		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-973-278-103 +	45.00	102.27	175.72		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-522-2 +	45.00	88.35	1.0e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-969-347-172 -	45.00	85.39	1.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-95 -	45.00	85.38	1.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-795-668-1 -	45.00	37.71	2.4e+05		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-834-975-773 +	44.50	95.36	426.44		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1830 -	44.50	91.29	718.70		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-353 -	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2902 -	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2903 -	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2904 -	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2905 +	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2906 +	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2907 +	44.50	71.93	8.5e+03		
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2908 +	44.50	71.93	8.5e+03		

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-969-708-319
seq_documentation_block:
; Sequence 319, Application US/09969708
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 319
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-319

alignment_scores:

Quality:	53.00	Length:	17
Ratio:	3.786	Gaps:	0
Percent Similarity:	82.353	Percent Identity:	52.941

alignment_block:

US-09-215-435-177_COPY_1_24 x US-09-969-708-319

Align seg 1/1 to: US-09-969-708-319 from: 1 to: 473

8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeuL 24
||||||| :|||||:|||||:| :|||||:|
15 TTTATTATTATTATTTTGTTCAGACAAACATTGATTTCTGG 64

24 Y 24
65 A 65

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-8488

seq_documentation_block:

; Sequence 8488, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.10130
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 8488
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-8488

alignment_scores:

Quality:	51.00	Length:	14
----------	-------	---------	----

us-09-215-435-177_copy_1_24.rnpn

Thu Nov 15 10:51:56 2001

```

Ratio: 4.250          Gaps: 0
Percent Similarity: 85.714    Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-8488  ..
Align seg 1/1 to: US-09-388-906A-8488 from: 1 to: 274
      8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 214

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-9659

seq_documentation_block:
; Sequence 9659, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.1013U
; CURRENT FILING DATE: 1999-09-01
; CURRENT APPLICATION NUMBER: US/09/388,906A
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9659
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-9659

alignment_scores:
Quality: 51.00          Length: 14
Ratio: 4.250           Gaps: 0
Percent Similarity: 85.714    Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-9659  ..
Align seg 1/1 to: US-09-388-906A-9659 from: 1 to: 333
      8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
178 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTG 219

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-11087

seq_documentation_block:
; Sequence 11087, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.1013U
; CURRENT FILING DATE: 1999-09-01
; CURRENT APPLICATION NUMBER: US/09/388,906A
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11087
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-11087

alignment_scores:
Quality: 51.00          Length: 14
Ratio: 4.250           Gaps: 0
Percent Similarity: 85.714    Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-22403  ..
Align seg 1/1 to: US-09-388-906A-22403 from: 1 to: 354
      8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTG 240

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-22094

seq_documentation_block:
; Sequence 22094, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.1013U
; CURRENT FILING DATE: 1999-09-01
; CURRENT APPLICATION NUMBER: US/09/388,906A
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22094
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-22094

alignment_scores:
Quality: 51.00          Length: 14
Ratio: 4.250           Gaps: 0
Percent Similarity: 85.714    Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-22094  ..
Align seg 1/1 to: US-09-388-906A-22094 from: 1 to: 386
      8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTG 282

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-22403

seq_documentation_block:
; Sequence 22403, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.1013U
; CURRENT FILING DATE: 1999-09-01
; CURRENT APPLICATION NUMBER: US/09/388,906A
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22403
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-22403

alignment_scores:
Quality: 51.00          Length: 14
Ratio: 4.250           Gaps: 0
Percent Similarity: 85.714    Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-22403  ..

```

Align seg 1/1 to: US-09-388-906A-22403 from: 1 to: 409

```

      8 PheLeuLeuArgPhePheLeuGlyPheCysSerLeuAsnThrLeu 21
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
    280 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 321

```

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-18975

```

seq_documentation_block:
; Sequence 18975, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; FILE REFERENCE: 11000.10130
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18975
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-18975

```

```

alignment_scores:
  Quality: 51.00      Length: 24
  Ratio: 3.188       Gaps: 0
  Percent Similarity: 66.667   Percent Identity: 45.833

```

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-18975/rev ..

Align seg 1/1 to reverse of: US-09-388-906A-18975 from: 1 to: 412

```

      1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCys 17
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
    408 CTCACCTCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 359
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
    358 TTGCAATAGCTTGTCTACTTGGGA 337

```

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-9295

```

seq_documentation_block:
; Sequence 9295, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9295
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1233)
US-09-815-242-9295

alignment_scores:
  Quality: 50.00      Length: 17
  Ratio: 3.571       Gaps: 0
  Percent Similarity: 82.353   Percent Identity: 47.059

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-815-242-9295 ..

Align seg 1/1 to: US-09-815-242-9295 from: 1 to: 1233

      7 GlyPheLeuLeuArgPhePheLeuPheCysSerLeuAsnThrLeuLeu 23
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
    1088 GGATTTTGAAGGATGGTGTTCGCTTTAAACAGAAATTTAAT 1137
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      23 u 23
      1138 G 1138

```

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-9530

```

seq_documentation_block:
; Sequence 9530, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9530
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:

```

us-09-215-435-177_copy_1_24.rnpn

Thu Nov 15 10:51:56 2001

```
; NAME/KEY: CDS
; LOCATION: (1)...(1233)
US-09-815-242-9530

alignment_scores:
  Quality: 50.00      Length: 17
  Ratio: 3.571       Gaps: 0
  Percent Similarity: 82.353   Percent Identity: 47.059

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-815-242-9530 ..
Align seg 1/1 to: US-09-815-242-9530 from: 1 to: 1233
7 GlyPheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeuLeu 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1088 GGATTTTGATGGAGTGATGGTGTGTTTGGGTTTAAACAGAGATTTTAAT 1137
23 u 23
1138 G 1138

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-214

seq_documentation_block:
; Sequence 214, Application US/099711392
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 3175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Template ID: 332263.1
; NAME/KEY: unsure
; LOCATION: 1854-1892
; OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-214

alignment_scores:
  Quality: 50.00      Length: 14
  Ratio: 4.167       Gaps: 0
  Percent Similarity: 85.714   Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-971-392-214 ..
Align seg 1/1 to: US-09-971-392-214 from: 1 to: 3175
8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
939 TTCAGATCAGATCTTCATTTTTCGAGCTACGATCTTCTT 980

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-8149

seq_documentation_block:
; Sequence 8149, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; TITLE OF INVENTION: Methods For Their Use
; FILE REFERENCE: 11000.10130
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8149
; LENGTH: 208
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-8149

alignment_scores:
  Quality: 49.00      Length: 14
  Ratio: 4.083       Gaps: 0
  Percent Similarity: 85.714   Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-8149 ..
Align seg 1/1 to: US-09-388-906A-8149 from: 1 to: 208
8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
111 TTTCCTCTCTCTTTTCTTTTCTTTTGTGTCGTCTCTTTG 152

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-10026

seq_documentation_block:
; Sequence 10026, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; TITLE OF INVENTION: Methods For Their Use
; FILE REFERENCE: 11000.10130
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10026
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-10026

alignment_scores:
  Quality: 49.00      Length: 14
  Ratio: 4.083       Gaps: 0
  Percent Similarity: 85.714   Percent Identity: 64.286

alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-10026 ..
Align seg 1/1 to: US-09-388-906A-10026 from: 1 to: 351
8 PheLeuLeuArgPhePheIlePheCysSerLeuAsnThrLeu 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 TTTCCTCTCTCTTTTCTTTTCTTTTGTGTCGTCTCTTTG 250

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-388-906A-5952

seq_documentation_block:
; Sequence 5952, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; TITLE OF INVENTION: Methods For Their Use
; FILE REFERENCE: 11000.10130
```

```
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5952
; LENGTH: 125
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-388-906A-5952
```

```
alignment_scores:
  Quality: 48.00      Length: 12
  Ratio: 4.364        Gaps: 0
Percent Similarity: 91.667 Percent Identity: 75.000
```

```
alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-388-906A-5952 ..
```

```
Align seg 1/1 to: US-09-388-906A-5952 from: 1 to: 125
```

```
6 LeuGlyPheLeuLeuArgPhePheLeuPheCysSer 17
|||||:  |||||||:|||||:|||||:|||||
22 CRAGGATTTCTCGGATCTCTCTCTCTCTCTCG 57
```

```
seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-4662
```

```
seq_documentation_block:
; Sequence 4662, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4662
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(435)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-4662
```

```
alignment_scores:
  Quality: 47.00      Length: 23
  Ratio: 2.611        Gaps: 0
Percent Similarity: 78.261 Percent Identity: 39.130
```

```
alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-922-340-4662/rev ..
```

```
Align seg 1/1 to reverse of: US-09-922-340-4662 from: 1 to: 435
```

```
1 MetLysSerAlaLysLeuGlyPheLeuLeuArgPhePheLeuPheCysSer 17
:::|||||: ||| |||||
234 CTTAGATCAAGTACTTAACCTTTTGAGCTTCAGTTTATCATCTTTGTAA 185

17 rLeuAsnThrLeuLeuLeu 23
:::|||||:|||||:
184 AATGGATACATCTGTATG 166
```

```
seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-111
```

```
seq_documentation_block:
; Sequence 111, Application US/09954456
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-111
```

```
alignment_scores:
  Quality: 47.00      Length: 15
  Ratio: 3.615        Gaps: 0
Percent Similarity: 86.667 Percent Identity: 60.000
```

```
alignment_block:
US-09-215-435-177_COPY_1_24 x US-09-954-456-111 ..
```

```
Align seg 1/1 to: US-09-954-456-111 from: 1 to: 572
```

```
9 LeuLeuArgPhePheLeuPheCysSerLeuAsnThrLeuLeuLeu 23
:::|||||:|||||:|||||:|||||:|||||
409 GTATTGCGACTGTTTCTTCTCTTAAATATTATTGCTT 453
```

us-09-215-435-177_copy_1_24.rnpn

Thu Nov 15 10:51:56 2001

Thu Nov 15 10:51:56 2001

```

seq_name: gb_gss3:CNS02GG5
seq_documentation_block:
LOCUS      CNS02GG5      980 bp      DNA      GSS      13-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
            137821 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL196286
VERSION    AL196286.1 GI:7834436
KEYWORDS   GSS; genome survey sequence.
SOURCE     Tetraodon nigroviridis.
ORGANISM   Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE  1 (bases 1 to 980)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 980)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
JOURNAL   Unpublished
REFERENCE  3 (bases 1 to 980)
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT   This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
            1..980
                /organism="Tetraodon nigroviridis"
                /db_xref="taxon:99883"
                /clone="137H21"
                /clone_lib="G"
                /note="Genoscope sequence ID : COAG137CD11L1P1-end : T7"
BASE COUNT 285 a 171 c 184 g 335 t 5 others
ORIGIN
alignment_scores:
Quality: 62.00 Length: 23
Ratio: 3.647 Gaps: 0
Percent Similarity: 73.913 Percent Identity: 52.174
alignment_block:
US-09-215-435-177_COPY_1_24 x CNS02GG5
Align seg 1/1 to: CNS02GG5 from: 1 to: 980
2 LysSerAlaLysLeuGlyPheLeuLeuArgPhePheIlePheCysSerLe 18
|||||
835 AAATGTCACCAACTATTTTATTATTAACAATTTTGTGCTGCTTT 884
18 uASThrLeuLeuLeuGly 24
885 CTCATCTGCGCTGTCGCG 903

```

```

seq_name: gb_gss32:AZ760839
seq_documentation_block:
LOCUS      AZ760839      487 bp      DNA      GSS      16-FEB-2001
DEFINITION 1M0554B19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0554B19 R, DNA sequence.
ACCESSION  AZ760839
VERSION    AZ760839.1 GI:12869114
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 487)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0554 row: B column: 19
            Seq primer: CACACAGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 487.
FEATURES   Location/Qualifiers
            1..487
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGC1M0554B19"
                /clone_lib="Mouse 10kb plasmid UUGC1M library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /note="Vector: pWD42nv; Purified genomic DNA from M.
            musculus C57BL/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a
            0.005 inch orifice at constant velocity. The sheared DNA
            was blunt end-repaired with T4 DNA polymerase and T4
            polynucleotide kinase. Adaptor oligonucleotides were
            ligated to the blunt ends in high molar excess. The
            adapted DNA was purified and size-selected for a 9.5 to
            10.5 kb range using preparative agarose gel
            electrophoresis. Vector DNA was prepared from a derivative
            of pWD42 (gil4732114|gb|AF129072.1), a copy-number
            inducible derivative of plasmid R1. The vector was ligated
            with adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adapted vector DNA, and transformed into
            chemically-competent E. coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
BASE COUNT 94 a 80 c 90 g 223 t
ORIGIN
alignment_scores:
Quality: 59.00 Length: 16
Ratio: 3.933 Gaps: 0
Percent Similarity: 93.750 Percent Identity: 62.500
alignment_block:
US-09-215-435-177_COPY_1_24 x AZ760839

```

Align seg 1/1 to: A2222099 from: 1 to: 510

 λ

```
/tab_host="DHIOB"
```

```

, /note="Site_1: Sali; Site_2: BamHI; cdna library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cdna was
primed with a primer [5'
GAGAGAGAAGGATCCAAAGACGCTTTTTTTTTTTTTTNN 3'], cdna was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper.. cdna went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cdna was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGATTAAATTAATTAATTCCTCCCCCCCC
3']. cdna was cloned into the XhoI and BamHI sites.
vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end:
BamHI
50 a 71 c 50 g 112 t
BASE COUNT
ORIGIN

alignment_scores:
    Quality: 58.50      Length: 21
    Ratio: 3.900        Gaps: 1
    Percent Similarity: 71.429    Percent Identity: 61.905

alignment_block:
US-09-215-435-177_COPY_1_24 x BB351590 ..
Align seg 1/1 to: BB351590 from: 1 to: 283

6 LeuGlypheLeuLeuArgPhePheLePheCys.....SerLeuAs 19
||||||| |||||||:|||||
14 TTGGGCTTACCCTTCGGTTTTTGTGTTTCTGGTCCCTGAA 63

19 nThrLeuLeu 23
|::|||
64 TCTGGCATTTCTG 76

seq_name: gb_gss11:AQ414393

seq_documentation_block:
LOCUS AQ414393 464 bp DNA GSS 23-MAR-1999
DEFINITION RPCI-11-169M20.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-169M20
, DNA sequence.
ACCESSION AQ414393
VERSION AQ414393.1 GI:4470175
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pietier de Jong
(pieteredejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
seq primer: SP6

```

```

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BS1-av1-e-09-0-UI"
/clone_lib="UI-R-BS1"
/dev_stages="embryonic 13 dpc"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-BS1
library is derived from 13 dpc whole embryo tissue. For a
detailed description of the library from which this clone
was derived, please visit our web site at
ratest.eng.uiowa.edu.
TAG_LIB=UI-R-BS1
TAG_TISSUE=embryo at 13 dpc
TAG_SEQ=AATCC"

```

```

BASE COUNT      74 a      40 c      83 g      94 t      1 others
ORIGIN

```

```

alignment_scores:
  Quality:      57.50      Length:      21
  Ratio:        3.382      Gaps:      1
Percent Similarity: 80.952      Percent Identity: 61.905

```

```

alignment_block:
US-09-215-435-177_COPY_1_24 x BE117416/rev ..

```

```

Align seg 1/1 to reverse of: BE117416 from: 1 to: 292

```

```

6 LeuGlyPheLeuLeuArgPheIlePheCys.....SerLeuAs 19
|||||
104 CTGGATTCTACTCGTTTCTTTTTCCTGGAAGCGAAGCTTAA 55
19 nThrLeuLeuLeu 23
|||||
54 TACTGTGATTCTA 42

```

```

seq_name: gb_gssl1:AQ395192

```

```

seq_documentation_block:
LOCUS      AQ395192      542 bp      DNA      GSS      06-MAR-1999
DEFINITION CITBI-E1-2556P5.TR CITBI-E1 Homo sapiens genomic clone 2556P5, DNA
sequence.
ACCESSION  AQ395192
VERSION    AQ395192.1  GI:4366218
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 542)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE      Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1997)
COMMENT    Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

```

```

FEATURES
            source
            1..542
            /organism="Homo sapiens"
            /db_xref="taxon:9606"

```

```

/clone="2556P5"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT      142 a      88 c      88 g      224 t
ORIGIN

```

```

alignment_scores:
  Quality:      57.50      Length:      25
  Ratio:        3.594      Gaps:      1
Percent Similarity: 64.000      Percent Identity: 48.000

```

```

alignment_block:
US-09-215-435-177_COPY_1_24 x AQ395192 ..

```

```

Align seg 1/1 to: AQ395192 from: 1 to: 542

```

```

7 GlyPheLeu.....LeuArgPhePheIlePheCy 16
|||||
144 GGTTCCTGTCCTCGCTTCAGATTTCCTTTGTATTTCTTTCTG 193
16 sSerLeuAsnThrLeuLeuLeuGly 24
|||||
194 TAGTTTAATACAAATATGCGTAGGG 218

```

```

seq_name: gb_gssl1:AQ393358

```

```

seq_documentation_block:
LOCUS      AQ393358      589 bp      DNA      GSS      06-MAR-1999
DEFINITION CITBI-E1-2556D1.TR CITBI-E1 Homo sapiens genomic clone 2556D1, DNA
sequence.
ACCESSION  AQ393358
VERSION    AQ393358.1  GI:4364381
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 589)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1997)
COMMENT    Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

```

```

FEATURES
            source
            1..589
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="2556D1"
            /clone_lib="CITBI-E1"
            /sex="male"
            /cell_type="sperm"
            /notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
            CalTech Human BAC Library D"
BASE COUNT      161 a      93 c      98 g      237 t
ORIGIN

```

二

```

|||||:::  |||||:::  |||||:::  |||||:::  |||||:::
174 CTCAATACCTTACTGTGATTGGCC 198
seq_name: gb_est100:BG399970

seq_documentation_block:
LOCUS      BG399970          903 bp    mRNA           EST       12-MAR-2001
DEFINITION 602442023F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4557893 5' ,
            mRNA sequence.
ACCESSION   BG399970
VERSION     BG399970
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 903)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            CDNA Library Preparation: CLONTECH Laboratories, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
Plate: LLCM1262 row: n column: 06
High quality sequence stop: 573.
Location/Qualifiers
source      1..903
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4557893"
            /clone_lib="NIH_MGC_75"
            /lab_host="DH10B (T1 phage-resistant)"
            /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
            SfiI (ggccgctcgccc); Site_2: SfiI (ggccattatgcc); 5' and
            3' adaptors were used in cloning as follows: 5' adaptor
            sequence: 5'-CAGCGCCATTATGCC-3' and 3' adaptor sequence:
            5'-ATTGAGAGCGCGAGCGCGCACATG-dt(30)BN-3' (where B = A,
            C, or G, and N = A, C, G, or T). Average insert size 1.65
            kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
            by PCR. This library was enriched for full-length clones
            and was constructed by Clontech Laboratories (Palo Alto,
            CA). Note: this is a NIH/MGC Library."
BASE COUNT 333 a 170 c 160 g 240 t
ORIGIN

alignment_scores:
Quality: 56.00 Length: 21
Ratio: 3.733 Gaps: 1
Percent Similarity: 71.429 Percent Identity: 57.143

alignment_block:
US-09-215-435-177_COPY_1_24 x BG399970/rev ..
Align seg 1/1 to reverse of: BG399970 from: 1 to: 903

      8 PheLeuLeuArgPhePhe.....IIlePheCysSerLeuAsnTh 20
        ||||| |||||
      179 TTTCCTTAGATTTCCTTTCTCTCTTGCTTCGTTTTCTGCAGTTTGATAT 130
              ::::::::::::::::::::

      20 rLeuLeuLeuGly 24
         ::::::|||||
      129 GATATTTCTAGGT 117

seq_name: gb_gss37:PT010M15R
seq_documentation_block:
LOCUS      AQ879006          1524 bp    DNA           GSS       09-NOV-1999
DEFINITION HS_3154_A2_E10_T7c CIT Approved Human Genomic sperm Library D Homo
            sapiens genomic clone plate=3154 Col=20 Row=1, DNA sequence.

```

```

CESSION      AQ879006
RSION        AQ879006.1   GI:6310473
URCE         GSS.
WORDS        HUMAN.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 1524)
AUTHORS      Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
              Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
              Hood,L.
TITLE        Sequence-tagged connectors: A sequence approach to mapping and
              scanning the human genome
JOURNAL       Proc..Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE       99380589
COMMENT       Contact: Mahairas GG, Wallace JC, Hood L
              High Throughput Sequencing Center
              University of Washington
              401 Queen Anne Avenue North, Seattle, WA 98109, USA
              Tel: (206) 616-3618
              Fax: (206) 616-3887
              Email: jwallace@u.washington.edu
              Clones may be purchased from Research Genetics (info@resgen.com).
              BAC end Web Server: http://www.htsc.washington.edu
              Plate: 3154 row: I column: 20
              Seq primer: T7
              Class: BAC ends
              High quality sequence stop: 1524.
FEATURES             location/Qualifiers
source               1..1524
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_plate=3154 Col=20 Row=I"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B."
BASE COUNT          327 a   327 c   324 g   458 t      88 others
ORIGIN
alignment_scores:
Quality:           56.00                Length:           18
Ratio:             4.000                 Gaps:              0
Percent Similarity: 77.778              Percent Identity:  61.111
alignment_block:
US-09-215-435-177_COPY_1_24 x AQ879006 ..
Align seg 1/1 to: AQ879006 from: 1 to: 1524
2   LyssrAlatysLeuglyPheLeuArgpheHelLePheCysSerLe 18
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
404 AAATCATGTTATTGGTGTTCCTTTTGTTGTTGTTGTTGTTGTCACG 453
18 uAsn 19
|||
454 CAAC 457
seq_name: gb_est98:BG287285
seq_documentation_block:
LOCUS      BG287285            1047 bp            mRNA                    EST
DEFINITION 60238223F1 NIH_MGC_93 Homo sapiens cdna clone IMAGE:4499909 5',
mRNA sequence.
ACCESSION  BG287285
VERSION    BG287285.1      GI:13040973
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 1047)

```


OM of: US-09-215-435-177_COPY25_99 to: GenEmbl:* out_format : pfs
Date: Nov 15, 2001 12:41 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O/cgml_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=DILOSUM62 -TRANS=human40.cdi
-LIST=45 -MODE=LOCAL -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -DOCAL=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=6
-MAXLEN=2000000000 -USER=US09215435 -CGML_1_11006 -NCPV=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-215-435-177_COPY25_99
Query length: 75
Database: GenEmbl:*
Database sequences: 1344157
Database length: -856060004
Search time (sec): 7698.940000

score_list:	Strd	Orig	Zscore	Escore	Len	Documentation
gb_pr9:HSDJ60101	+	341.00	675.46	2.2e-29	33458	AL109656 Human DNA sequence f
gb_hg9:AC023971	+	341.00	662.63	1.1e-28	153788	AC023971 Homo sapiens chromo
gb_om:BTU35642	+	134.00	268.46	1.0e-06	1198	U35642 Bos taurus alpha-micro
gb_in3:CEIC37C3	+	133.00	236.34	6.3e-05	42521	U64857 Caenorhabditis elegans
gb_hg1:AC006915	+	133.00	226.68	0.0002	134095	AC006915 Caenorhabditis eleg
gb_hg6:AC031993	+	131.00	231.32	0.0001	46847	AC019933 Drosophila melanoga
gb_in1:AC005149	+	131.00	226.75	0.0002	80689	AC030353 Drosophila melanoga
gb_in1:AC003579	+	131.00	225.04	0.0003	98942	AC005149 Drosophila melanoga
gb_om:MEELACP	+	130.00	268.88	9.7e-07	304383	AC0003579 Drosophila melanoga
gb_pr2:163359	+	130.00	258.42	3.7e-06	419	AJ000490 Macropus eugenii mRNA
gb_pr9:HSABMBG3	+	129.50	247.64	1.5e-05	1455	I63559 Sequence 6 from patent
gb_hg19:AL137850	+	129.50	246.30	1.8e-05	4627	I54818 Human gene for alpha-1
gb_hg3:AC011161	+	129.50	216.22	0.0008	5423	M88249 Human inter-alpha-tryps
gb_sy:SYNB0VTRY1	+	129.50	215.94	0.0009	193939	AL137850 Homo sapiens chromo
gb_in2:AF078161	+	129.00	239.30	4.3e-05	200532	AC011161 Homo sapiens clone
gb_in2:CBRG45N02	+	129.00	227.89	0.0002	301	M31335 Bovine synthetic colost
gb_pr8:HSALM1CR	+	128.50	263.37	2.0e-06	11008	AF078161 Manduca sexta lacuna
gb_pr2:105217	+	128.50	256.74	4.6e-06	42726	AC084649 Caenorhabditis brig
gb_pr9:HSBHR	+	128.50	256.66	4.6e-06	555	E12656 Synthetic DNA encoding u
gb_pat1:A31024	+	127.00	269.19	9.3e-07	12321	X04494 Human mRNA for alpha-1
gb_pat1:A19210	+	127.00	269.19	9.3e-07	1239	I05217 Sequence 7 from patent
gb_pat1:A19212	+	127.00	262.60	2.2e-06	1339	X04225 Human mRNA for protein
gb_pat2:129002	+	127.00	262.60	2.2e-06	191	A31024 DNA fragment for bikunin
gb_om:RVU34208	+	126.00	261.35	2.5e-06	418	A31025 DNA fragment for bikunin
gb_pat2:EL3092	+	124.00	255.96	5.1e-06	418	A19212 Human kunitz type protea
gb_rol:MMAMB	+	124.00	247.57	1.5e-05	418	I29002 Sequence 17 from patent
gb_rol:MUSAM1ATI	+	124.00	247.19	1.6e-05	378	U34208 Trichosurus vulpecula ea
gb_om:OAU00163	+	123.50	262.62	2.2e-06	1234	D28812 Mouse mRNA for alpha-1
gb_om:AF241828	+	123.50	247.12	1.6e-05	97385	AL050348 Human DNA sequence f
gb_pat1:129001	+	123.00	254.19	6.4e-06	174	U00163 Ovis aries secretory pro
gb_in3:CEW01F3	+	123.00	254.19	6.4e-06	1098	AF241828 Ovis aries kunitz dom
gb_hg24:CEV44A6	+	123.00	198.19	0.0084	418	A19210 Human kunitz type protea
gb_rol:S87544	+	122.00	243.50	2.5e-05	3797	I292815 Caenorhabditis elegans
gb_om:AF008648	+	122.00	240.51	3.7e-05	326074	I298863 Caenorhabditis elegans
gb_pr9:AK024817	+	121.00	245.33	2.0e-05	1162	S87544 polyprotein l-microglob
gb_om:SY61902	+	121.00	243.20	2.6e-05	1657	AF008648 Cyprinus carpio unknd
					728	AK024817 Homo sapiens cDNA: FLJ
					938	S61902 tissue factor pathway in

gb_pat2:IL14875 + 121.00 242.84 2.7e-05 979 ! IL14875 Sequence 1 from paten
gb_pat2:I92685 + 121.00 242.84 2.7e-05 979 ! I92685 Sequence 1 from paten
gb_pr10:HUMHTFP + 121.00 242.84 2.7e-05 979 ! L27624 Homo sapiens tissue f
gb_om:OCLIPOR + 121.00 242.17 3.0e-05 1060 ! X54708 Rabbit mRNA for lipo

seq_name: gb_pr9:HSDJ60101

seq_documentation_block:
LOCUS HSDJ60101 33458 bp DNA PRI 04-APR-2001
DEFINITION Human DNA sequence from clone RP4-60101 on chromosome 20. Contains
the gene for a novel protein with Kunitz/Bovine pancreatic trypsin
inhibitor domain, the 5' end of a novel gene, part of the gene for
a novel protein with WAP-type (Whey Acidic Protein) 'four-disulfide
core' domain, ESTs and GSSs, complete sequence.

ACCESSION AL109656
VERSION AL109656.10 GI:10944958
KEYWORDS HTG; Kunitz; WAP.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Submitted (03-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
TITLE CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
JOURNAL requests: clonerequest@sanger.ac.uk
1 (bases 1 to 33458)

COMMENT
On Oct 21, 2000 this sequence version replaced gi:9795206.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RP4-60101 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP3-447F3 is at 33359 in this sequence.
The true right end of clone RP4-688G8 is at 100 in this sequence.
RP4-60101 is from the library RPCI-4 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.

FEATURES
Location/Qualifiers
1..33458
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone_lib="RP4-60101"
/clone="RPCI-4"
3..225
/note="LIM4 repeat: matches 3977..4203 of consensus"
repeat_region
225..428
/note="LIM4 repeat: matches 4321..4525 of consensus"
repeat_region
429..696
/note="AluY repeat: matches 2..302 of consensus"
repeat_region
697..778

repeat_region	/note="L1M4 repeat: matches 4525. .4606 of consensus" 779. .1089	repeat_region	/note="L2 repeat: matches 2666. .2732 of consensus" 13202. .13331
repeat_region	/note="AluSx repeat: matches 1. .311 of consensus" 1090. .1146	repeat_region	/note="AluSg/x repeat: matches 183. .312 of consensus" 13332. .15510
repeat_region	/note="L1M4 repeat: matches 4606. .4662 of consensus" 1165. .1293	repeat_region	/note="L1MB1 repeat: matches 3916. .6171 of consensus" 15534. .15832
repeat_region	/note="L1M4 repeat: matches 1. .134 of consensus" 1351. .1730	repeat_region	/note="AluSg repeat: matches 1. .299 of consensus" 15846. .16746
repeat_region	/note="L1MB5 repeat: matches 5005. .5402 of consensus" 1731. .2032	repeat_region	/note="L1M4 repeat: matches 3472. .4394 of consensus" 16739. .18594
repeat_region	/note="AluSx repeat: matches 1. .301 of consensus" 2056. .2547	repeat_region	/note="L1PA11 repeat: matches 11. .1921 of consensus" 18595. .18888
repeat_region	/note="L1MB5 repeat: matches 5374. .5865 of consensus" 2543. .2729	repeat_region	/note="AluSg repeat: matches 1. .295 of consensus" 18889. .20972
repeat_region	/note="L1MC/D repeat: matches 5428. .5609 of consensus" 2730. .3022	repeat_region	/note="L1PA11 repeat: matches 1. .310 of consensus" 20973. .21280
repeat_region	/note="AluSg repeat: matches 1. .293 of consensus" 3023. .3181	repeat_region	/note="AluSg repeat: matches 1. .310 of consensus" 21281. .23911
repeat_region	/note="L1MC/D repeat: matches 5609. .5769 of consensus" 3361. .3895	repeat_region	/note="L1M4 repeat: matches 4390. .4613 of consensus" 23926. .24129
repeat_region	/note="L1 repeat: matches 3599. .4167 of consensus" 3986. .4281	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" 24130. .24438
repeat_region	/note="L2 repeat: matches 534. .834 of consensus" 4526. .4828	repeat_region	/note="AluSx repeat: matches 1. .310 of consensus" 24439. .24513
misc_feature	/note="AluSx repeat: matches 1. .300 of consensus" 5372. .5726	repeat_region	/note="L1M4 repeat: matches 4613. .4683 of consensus" 24790. .25008
repeat_region	/note="match: GSS: Em:AQ076126" 5394. .5470	repeat_region	/note="MIR repeat: matches 1. .242 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
gene	/note="MIR repeat: matches 48. .142 of consensus" 7401. .7609	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
CDS	/gene="dj60101.3" <7401. .>7609	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
	/note="possibly a pseudogene Tr:044131 Wp:C08G9"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
	/evidence=not_experimental /product="dj60101.3 (novel protein with WAP-type (Whey Acidic Protein) 'four-disulfide core' domain)"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
	/db_xref="GI:13559237" /translation="FPGSSGLSSEHYSLLVRPFSYLCVPEVILEPPPCISAPEN CTHCTMQEDCEKGFQCCSFCGIV"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
misc_feature	/complement(7623. .7780) /gene="dj688G8.4"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="Continues in Em:AL031671 as dj688G8.4 match: ESTs: Em:AI808149 Sw:P00993 Sw:Q28631 Sw:P14730 Tr:044131 Wp:C08G9"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/evidence=not_experimental /product="dj60101.2 (a putative novel protein)"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/complement(7623. .7780) /gene="dj688G8.4"	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="L2 repeat: matches 2596. .2670 of consensus" 8352. .8422	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="MIR repeat: matches 9. .260 of consensus" 8434. .8656	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	9290. .9419	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
misc_feature	/note="MIR repeat: matches 47. .183 of consensus" complement(9433. .9899)	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="match: GSS: Em:AQ476416" 9538. .9830	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
misc_feature	/note="AluSg repeat: matches 1. .293 of consensus" 9923. .10651	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="match: GSS: Em:AQ478366" 10453. .10647	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
misc_feature	/note="MER3 repeat: matches 1. .204 of consensus" complement(10590. .10982)	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="match: GSS: Em:AQ223524" 11007. .11080	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="MER69 repeat: matches 2. .77 of consensus" 11529. .11954	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)
repeat_region	/note="MER69 repeat: matches 2070. .2511 of consensus" 13092. .13158	repeat_region	/note="L1M4 repeat: matches 1. .310 of consensus" Join(25397. .25528,26926. .27103,28675. .28876)

alignment_scores:
Quality: 341.00 Length: 60
Ratio: 5.683 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-177_COPY_25_99 x HSDJ60101 ..
Align seg 1/1 to: HSDJ60101 from: 1 to: 33458

15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
 |||||
 26925 GATCCCTGCAATTTGGACATGAATTTTGGAAAGCTGCTATGAAGTTCACCT 26974
 |||||
 31 eAqTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
 |||||
 26975 TAGATATTTCTACACAGAACCTCCAAAAGATGTGAACACTTTTGTCTCT 27024
 |||||
 48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
 |||||
 27025 CCGCGTGTATGCACACCTTAACAACTTCAAGCTTAAATAGACGTGAA 27074
 |||||
 65 ValAlaCysValAlaLysTyrLysProPro 74
 |||||
 27075 GTAGCCTGTGTGCAAAATACAAACCAACCG 27104
 |||||

seq_name: gb_htg9.AC023971

seq_documentation_block:
 LOCUS AC023971 153788 bp DNA HTG 07-APR-2000
 DEFINITION Homo sapiens chromosome 20 clone RP11-770N8 map 20, WORKING DRAFT
 SEQUENCE, 13 unordered pieces.
 ACCESSION AC023971
 VERSION AC023971.2 GI:7523758
 KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 153788)
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 20, clone RP11-770N8
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 153788)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collumore,A., Cooke,P.,
 DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
 Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
 Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
 McEvan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneses,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M.,
 Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
 Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
 Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
 Wu,X., Wyman,D., Ye.W.J., Young,G., Zainoun,J., Zimmer,A. and
 Zody,M.

Direct Submission
 Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 7, 2000 this sequence version replaced gi:7008875.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L6573
 Center clone name: 770_N8
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
 Consensus bases at least Q40
 Consensus quality: 148490 bases at least Q30
 Consensus quality: 150621 bases at least Q20
 Insert size: 150000; agarose-fp
 Insert size: 152588; sum-of-contigs
 Quality coverage: 5.2 in Q20 bases; agarose-fp
 Quality coverage: 5.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1333: contig of 1333 bp in length
 * 1334 1433: gap of 100 bp
 * 1434 3969: contig of 2536 bp in length
 * 3970 4069: gap of 100 bp
 * 4070 7211: contig of 3142 bp in length
 * 7212 7311: gap of 100 bp
 * 7312 13518: contig of 6207 bp in length
 * 13519 13618: gap of 100 bp
 * 13619 20876: contig of 7258 bp in length
 * 20877 20976: gap of 100 bp
 * 20977 28838: contig of 7862 bp in length
 * 28839 28938: gap of 100 bp
 * 28939 37590: contig of 8652 bp in length
 * 37591 37690: gap of 100 bp
 * 37691 45814: contig of 8124 bp in length
 * 45815 45914: gap of 100 bp
 * 45915 62601: contig of 16687 bp in length
 * 62602 62701: gap of 100 bp
 * 62702 81817: contig of 19116 bp in length
 * 81818 81917: gap of 100 bp
 * 81918 100273: contig of 18356 bp in length
 * 100274 100373: gap of 100 bp
 * 100374 125266: contig of 24893 bp in length
 * 125267 125366: gap of 100 bp
 * 125367 153788: contig of 28422 bp in length.

FEATURES

Location/Qualifiers
 source
 1..153788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20"
 /clone="RP11-770N8"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1..1333
 /note="assembly_fragment"
 misc_feature
 1434..3969
 /note="assembly_fragment"
 misc_feature
 4070..7211
 /note="assembly_fragment"
 misc_feature
 7312..13518
 /note="assembly_fragment"
 misc_feature
 13619..20876
 /note="assembly_fragment"
 misc_feature
 20977..28838
 /note="assembly_fragment"
 misc_feature
 28939..37590
 /note="assembly_fragment"
 misc_feature
 37691..45814
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 45915..62601
 /note="assembly_fragment"
 misc_feature
 62702..81817
 /note="assembly_fragment"
 clone_end:T7

bikunin"
/codon_start=1
/product="alpha1-microglobulin/bikunin"
/protein_id="AAB07599.1"
/db_xref="GI:1016298"

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 42521)
Waterston, R.
Direct Submission
Submitted (08-JUL-1998) Department of Genetics, Washington

COMMENT University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
Submitted by: Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1HQ, England
e-mail: rw@nematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is 2K742, 200 bp overlap; 3' cosmid is F40A3, 200 bp
overlap. Actual start of this cosmid is at base position 1 of
CEL3C7C3; actual end is at 5024 of CELF40A3
NOTES:

Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in preparation).

FEATURES
Source

```
1. 42521
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/clone="C37C3"
/chromosome="V"
513..15440
/gene="C37C3.6b"
513..6558
/gene="C37C3.6a"
join(513..608,664..878,1380..1535,1583..1734,1785..2541,
2666..2897,2976..3097,3149..4733,4801..5584,5659..5871,
6194..6558)
/gene="C37C3.6a"
/notes="similar to the DPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor;
coded for by C. elegans cdna yk84d8.5; coded for by C.
elegans cdna yk117d1.5; coded for by C. elegans cdna
yk88e5.3; coded for by C. elegans cdna yk132d4.3; coded
for by C. elegans cdna yk123e8.3; coded for by C. elegans
cdna yk88e5.5; coded for by C. elegans cdna yk22b2.5;
coded for by C. elegans cdna yk132d4.5; coded for by C.
elegans cdna yk123e8.5; coded for by C. elegans cdna
CEESD32F; coded for by C. elegans cdna CEESH31FC; coded
for by C. elegans cdna CEESL67F; coded for by C. elegans
cdna CEESL73F; coded for by C. elegans cdna CEESP33F;
coded for by C. elegans cdna CEESP94F; coded for by C.
elegans cdna CESAC89F; coded for by C. elegans cdna
yk10f9.3; coded for by C. elegans cdna yk10f9.5; coded for
by C. elegans cdna yk22b2.3; coded for by C. elegans cdna
yk3e8.3; coded for by C. elegans cdna yk3e8.5; coded for
by C. elegans cdna yk41h8.5; coded for by C. elegans cdna
yk5c8.5; coded for by C. elegans cdna yk148d2.5; coded for
by C. elegans cdna yk152f10.5"
/codon_start=1
/protein_id="AAC25867.1"
/db_xref="GI:3294501"
/translation="MRLLFSAALLLCVPTWAFSLSFSGSDVAQKPYLHPNSPPER
DPASSRMKRAYQYVVDGVSVDKSGQKETGNWGPWVPENECSRSGGGVQLEKQ
```

gene

gene

CDS

```
CSGDTGASVRYISCNLNACESGDTFRAEQCKSFNDALDGNVHKWTPYKGNKNCIELV
CKPESGNYFKWADKVVVDGDKSKNDICVDGECPLVPGCDGKLSGLKFKDCKGCDG
DGSTCKTIEGRFEDNRLSPGVHDIILKPEGATNIIKIOEARKSTNNLAKFKDCKGCDG
GNGLIOVEKEVEVGSTIFVYDDAEPETLSAOGPLSEELTALLFRKSRDFAIKVEFS
IPLEEVDMYKFDNWTFCVSCGKGVQTRNLICIDGKNKGRVEDDICEENNAIKPEF
EKSCETVDCAEWETIGDWESCTCGDGGQGVYVYCHQVFANGRRVTVDEGNCVTVER
PPVKQTCNRFACPEWQAPWMSACEKGDFAQYRSVTCSEKEGEKGLLAADACPAD
EQEKTDERTCNLPGCEGLTFTVGEMNLTCNDTEETRETKDSQGRAYFLERKCLV
DNSTIEPTDRSCATOPPEYETVSWSEKCTTECGHKTTRRVICAIHONGLEVPVE
EGHCQAEKPEKTNCTNEEKTGTWYTSWSECTAECGGSSQDRAVAVCLNVDKRPVPE
WCDAAVPSEKQDCNVDDCPVDSERGCCPDNSTFATGFENFCSNCSERFEGCCAD
NVTVATGPNKSGCEEFVESPUNLEADVANADAEASGDAPELCSVTNNEGEAVDECAT
IPATALLGDELIGENDTASNETIHCSKTEFGCCPDWYTAASGKNEGCSFVLGGC
NETQFGCHDDVTILARGANLEGCEPSCAASLYCKCKDKRTIAFGPHYSGERSFPC
ELSDFGCCPDGETAALGKNGTCGCECLTTKFGCCPDGKTRAGENKEGCPQYTRYGCCPDGE
CCPDGKSVAGKAGFYGCPESCAOSOFCCPDGKTRAGENKEGCPQYTRYGCCPDGE
TTALGPRNDGCDNCRYAKHGCCPDGETKALGPDGAGCPTTTPFLMGGTVAHPKIAA
CNQFQESGTGCGAGYKLAHYDTTEGRCNQFYGGCGGNDFNNFASQDMCTICVEPPG
KGRCLPRVDGDLRCDQLQPRYYLTHDSKHCVAFWWRGCLGNANFNFECSMFCKD
VGPYDAPTTAAPPPPPQNAQYLLTPEVQIEQISAEQPOQPOQPOQPOQPOQPOQ
VGRYDLPTAAPPQPPQNAQYLLTPEVQIEQISAEQPOQPOQPOQPOQPOQPOQ
PROSMEDICRSRODAPCETYSDOWFYNAFSECECTFTYGGCGNLNFRFRSKDCEOR
CFVHGAQPSAAROEQAQPAQPAQPSNIVSPPOOSASPVVPSNKSORDACHLN
VDQCRKGFADSWIYEVATGSCVTFYKTYGCGGNANRRASKDQCESLVCVKPASEASAG
IDGAAGINSVCDEAKDTPCTNFVTKWYTNKADGTGNFRFTYGGCGGNTRNFATIECOQ
AACQNHKDACOLPKVQGPCSGKHSYYNTTASHOCTFTYGGCGGNTRNFATIECOQ
RCPSKF"
join(513..608,664..878,1380..1535,1583..1734,1785..2541,
2666..2897,2976..3097,3149..4733,4801..5584,5659..5871,
6194..6547,7038..7134,7881..8009,8060..8410,8493..8571,
13650..13741,14139..14727,14940..15440)
/gene="C37C3.6b"
/notes="similar to the BPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor
(TFPI); coded for by C. elegans cdna yk257g8.5; coded for
by C. elegans cdna yk257g8.3; coded for by C. elegans cdna
yk185b12.3"
/codon_start=1
/protein_id="AAC25868.1"
/db_xref="GI:3294502"
/translation="MRLLFSAALLLCVPTWAFSLSFSGSDVAQKPYLHPNSPPER
DPASSRMKRAYQYVVDGVSVDKSGQKETGNWGPWVPENECSRSGGGVQLEKQ
CSGDTGASVRYISCNLNACESGDTFRAEQCKSFNDALDGNVHKWTPYKGNKNCIELV
CKPESGNYFKWADKVVVDGDKSKNDICVDGECPLVPGCDGKLSGLKFKDCKGCDG
DGSTCKTIEGRFEDNRLSPGVHDIILKPEGATNIIKIOEARKSTNNLAKFKDCKGCDG
GNGLIOVEKEVEVGSTIFVYDDAEPETLSAOGPLSEELTALLFRKSRDFAIKVEFS
IPLEEVDMYKFDNWTFCVSCGKGVQTRNLICIDGKNKGRVEDDICEENNAIKPEF
EKSCETVDCAEWETIGDWESCTCGDGGQGVYVYCHQVFANGRRVTVDEGNCVTVER
PPVKQTCNRFACPEWQAPWMSACEKGDFAQYRSVTCSEKEGEKGLLAADACPAD
EQEKTDERTCNLPGCEGLTFTVGEMNLTCNDTEETRETKDSQGRAYFLERKCLV
DNSTIEPTDRSCATOPPEYETVSWSEKCTTECGHKTTRRVICAIHONGLEVPVE
EGHCQAEKPEKTNCTNEEKTGTWYTSWSECTAECGGSSQDRAVAVCLNVDKRPVPE
WCDAAVPSEKQDCNVDDCPVDSERGCCPDNSTFATGFENFCSNCSERFEGCCAD
NVTVATGPNKSGCEEFVESPUNLEADVANADAEASGDAPELCSVTNNEGEAVDECAT
IPATALLGDELIGENDTASNETIHCSKTEFGCCPDWYTAASGKNEGCPSTLGGC
NETQFGCHDDVTILARGANLEGCEPSCAASLYCKCKDKRTIAFGPHYSGERSFPC
ELSDFGCCPDGETAALGKNGTCGCECLTTKFGCCPDGKTRAGENKEGCPQYTRYGCCPDGE
CCPDGKSVAGKAGFYGCPESCAOSOFCCPDGKTRAGENKEGCPQYTRYGCCPDGE
TTALGPRNDGCDNCRYAKHGCCPDGETKALGPDGAGCPTTTPFLMGGTVAHPKIAA
CNQFQESGTGCGAGYKLAHYDTTEGRCNQFYGGCGGNDFNNFASQDMCTICVEPPG
KGRCLPRVDGDLRCDQLQPRYYLTHDSKHCVAFWWRGCLGNANFNFECSMFCKD
VGPYDAPTTAAPPPPPQNAQYLLTPEVQIEQISAEQPOQPOQPOQPOQPOQPOQ
VGRYDLPTAAPPQPPQNAQYLLTPEVQIEQISAEQPOQPOQPOQPOQPOQPOQ
PROSMEDICRSRODAPCETYSDOWFYNAFSECECTFTYGGCGNLNFRFRSKDCEOR
CFVHGAQPSAAROEQAQPAQPAQPSNIVSPPOOSASPVVPSNKSORDACHLN
VDQCRKGFADSWIYEVATGSCVTFYKTYGCGGNANRRASKDQCESLVCVKPASEASAG
IDGAAGINSVCDEAKDTPCTNFVTKWYTNKADGTGNFRFTYGGCGGNTRNFATIECOQ
AACQNHKDACOLPKVQGPCSGKHSYYNTTASHOCTFTYGGCGGNTRNFATIECOQ
RCPSKF"
join(513..608,664..878,1380..1535,1583..1734,1785..2541,
2666..2897,2976..3097,3149..4733,4801..5584,5659..5871,
6194..6547,7038..7134,7881..8009,8060..8410,8493..8571,
13650..13741,14139..14727,14940..15440)
/gene="C37C3.6b"
/notes="similar to the BPTI/Kunitz family of inhibitors;
most similar to tissue factor pathway inhibitor precursor
(TFPI); coded for by C. elegans cdna yk257g8.5; coded for
by C. elegans cdna yk257g8.3; coded for by C. elegans cdna
yk185b12.3"
/codon_start=1
/protein_id="AAC25868.1"
/db_xref="GI:3294502"
/translation="MRLLFSAALLLCVPTWAFSLSFSGSDVAQKPYLHPNSPPER
DPASSRMKRAYQYVVDGVSVDKSGQKETGNWGPWVPENECSRSGGGVQLEKQ
```


Ratio: 3.639 Gaps: 0
Percent Similarity: 64.286 Percent Identity: 41.071
alignment_block:
US-09-215-435-177_COPY_25_99 x AC019993/rev ..
Align seg 1/1 to reverse of: AC019993 from: 1 to: 46847
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
5598 GAGGACTGTCATCAGCCCAAGAACCGTCTGTCGCTCTGTCTA 5549
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
5548 TCGCTACGCGCTCAACACGTGGATACACAATCCTCGGAGGAGTTCGTTACG 5499
48 erGlyCysAsnGlyAsnLeuAsnPheLysLeuLysIleGluArgGlu 64
5498 GTGGATGTCGCCGCAACAGAACATCTCGAATCCAGGAGCAGTCGCGAA 5449
65 ValAlaCysValAlaLys 70
5448 CAGGATGTTTGGTAAAG 5431
seq_name: gb_inl:AC003053
seq_documentation_block:
LOCUS AC003053 80689 bp DNA INV 17-JUL-1998
DEFINITION Drosophila melanogaster DNA sequence (P1 DS00216 (D160)), complete
sequence.
ACCESSION AC003053 AC002863 AC003283 AC003284 AC003285 AC003286 AC003287
AC002864 AC003288 AC003289 AC003290 AC003291 AC003292 AC003293
AC003294 AC003295 AC003296 AC003297
AC003053.1 GI:2584830
VERSION AC003053.1
KEYWORDS HTG.
SOURCE Drosophila melanogaster (Subclones in tet from P1 clone DS00216 (D160)) DNA.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera.
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 80689)
Celniker,S.E., George,R.A., Galle,R.F., Hoskins,R.A.,
Svirskas,R.R., Harris,N.L., Aghavani,A., Arcaina,T.T., Baxter,E.,
Blazej,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E.,
Flanagan,J., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,S.H., Lee,B., Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacieb,J.M., Park,S.,
Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitelaw,K.R.,
Yee,A., Zhang,R., Zieran,L.L. and Kimmel,B.
Sequencing of Drosophila chromosome 2L, region 24A3-24C2
Unpublished (1997)
2 (bases 1 to 80689)
Celniker,S.E., Aghavani,A., Arcaina,T.T., Baxter,E., Doyle,C.M.,
Farfan,D.E., Flanagan,J., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Kim,S.H., Ko,C.L., Li,M., Lomotan,M.A., Mazda,P.,
Mok,M.S., Nixon,K., Pacieb,J.M., Park,S., Pfeiffer,B., Punch,D.,
Santos,R.F., Snir,E., Stekov,V., Subramanian,S., Towne,B.,
Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.L. and
Kimmel,B.
Direct Submission
Submitted (04-NOV-1997) Berkeley Drosophila Genome Project, MS
74-157, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://fruitfly.berkeley.edu/sequence/>) or send
email to drosophila@hgsc.lbl.gov.
Library location: 24-3.

FEATURES
source
1..80689
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
/map="24A3-24C2"
/clone="P1 DS00216 (D160)"
/note="This sequence has not changed since its original
submission on 11/04/1997. It was resubmitted in order to
include all secondary accession numbers for the subclones
belonging to this clone."
BASE COUNT 23494 a 16620 c 16241 g 24334 t
ORIGIN
alignment_scores:
Quality: 131.00 Length: 56
Ratio: 3.639 Gaps: 0
Percent Similarity: 64.286 Percent Identity: 41.071
alignment_block:
US-09-215-435-177_COPY_25_99 x AC003053 ..
Align seg 1/1 to: AC003053 from: 1 to: 80689
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
6096 GAGGACTGTCATCAGCCCAAGAACCGTCTGTCGCTCTGTCTA 6145
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
6146 TCGCTACGCGCTCAACACGTGGATACACAATCCTCGGAGGAGTTCGTTACG 6195
48 erGlyCysAsnGlyAsnLeuAsnPheLysLeuLysIleGluArgGlu 64
6196 GTGGATGTCGCCGCAACAGAACATCTCGAATCCAGGAGCAGTCGCGAA 6245
65 ValAlaCysValAlaLys 70
6246 CAGGATGTTTGGTAAAG 6263
seq_name: gb_inl:AC005149
seq_documentation_block:
LOCUS AC005149 98942 bp DNA INV 20-JUN-1998
DEFINITION Drosophila melanogaster DNA sequence (P1S DS03364 (D237) and
DS02911 (D145)), complete sequence.
ACCESSION AC005149 AC004339 AC004278 AC003204 AC003205 AC002788 AC002789
AC002790 AC002791 AC002792 AC002793 AC002794 AC002795 AC002796
AC003206 AC002797 AC002798 AC002799 AC003207 AC002800 AC002801
AC002802 AC003208
AC005149.1 GI:3242679
VERSION AC005149.1
KEYWORDS HTG.
SOURCE Drosophila melanogaster (Subclones in tet from P1 clones DS03364 (D237) and DS02911 (D145)) DNA.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera.
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 98942)
Celniker,S.E., Aghavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
Lomotan,M.A., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M.,
Nixon,K., Pacieb,J.M., Park,S., Pfeiffer,B., Punch,D., Snir,E.,
Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.L.
and Kimmel,B.
Sequencing of Drosophila chromosome 2L, region 24A1-24C5
Unpublished (1997)
2 (bases 1 to 98942)
Celniker,S.E., Aghavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
Lomotan,M.A., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M.,
Nixon,K., Pacieb,J.M., Park,S., Pfeiffer,B., Punch,D., Snir,E.,
Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.L.
and Kimmel,B.
TITLE Sequencing of Drosophila chromosome 2L, region 24A1-24C5
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 98942)
AUTHORS Celniker,S.E., Aghavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,

Thu Nov 15 10:51:57 2001

```

Ananatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F.,
George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N.,
Sutton, G. G., Rogers, J. R., Tandel, M. D., Zhang, Q., Chen, L. X.,
Brandon, R. C., Rogers, Y. H., Blazer, R. G., Champe, M., Pfeiffer, B. D.,
Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor
Miklos, G. L., Abril, J. F., Agbayani, A., An, H. J.,
Andrews-Pfannkuch, C., Baldwin, D., Ballew, R. M., Basu, A.,
Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y.,
Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D.,
Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burkis, K. C.,
Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I.,
Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de
Pablos, B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M.,
Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C.,
Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S.,
Fleischmann, W., Foster, C., Gabrielian, A. E., Garg, N. S.,
Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z.,
Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J.,
Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J.,
Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z.,
Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C.,
Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A.,
Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C.,
McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C.,
Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L.,
Muzny, D. M., Nelson, D. L., Nelson, D. R., Pittman, G. S., Pan, S.,
Nusskern, D. R., Pacle, J. M., Palazzolo, M., Reinert, K., Remington, K.,
Pollard, J., Puri, V., Reese, M. G., Shen, H., Shue, B. C., Siden-Kiamos, I.,
Saunders, R. D., Scheeler, F., Smith, T., Spier, E., Spradling, A. C.,
Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R.,
Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A.,
Weinstock, G. M., Weissbach, J., Williams, S. M., Woodage, T.,
Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F.,
Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H.,
Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O.,
Gibbs, R. A., Myers, E. W., Rubin, G. M. and Venter, J. C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
20196006
2 (bases 1 to 304383)
Adams, M. D., Celniker, S. E., Gibbs, R. A., Rubin, G. M. and Venter, C. J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7295765.
Location/Qualifiers
1. 304383
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
/join(14340..14462,14669..15409,15464..15868)
/gene="CG2772"
/product="CT9119"
/db_xref="FLYBASE:FBan0002772"
/db_xref="FLYBASE:FBgn0031533"
<14340..>15868
/gene="CG2772"
/map="23F1-23F3"
/db_xref="FLYBASE:FBan0002772"
/db_xref="FLYBASE:FBgn0031533"
/join(14358..14462,14669..15409,15464..15868)
/gene="CG2772"
/note="CG2772 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0002772"
/db_xref="FLYBASE:FBgn0031533"
/protein_id="AAF51110.1"
/db_xref="GI:7295808"
/translaton="MLARVLTILAVASAGAADDPFDIDIPKRLKTSATRIAEHGY
PAESHFVETPDGYVLNVERIPKLSNGEGSEASRPVLIWHGLSCDCLLN
GPDALPYNDAGYDVLWGNAGNIYRNRLNVRHVPFKWESNHEISDLPATI
DYLAETGQQLHYVGHSGQCTSFVMSYRPEYNKIKTAHMLAPPVYMGNSTEGLI"

```

```

Lomontan, M. A., Mazda, P., Mok, M. S., Moshrefi, A. R., Moshrefi, M.,
Nixon, K., Pacle, J. M., Park, S., Pfeiffer, B., Punch, D., Sair, E.,
Twomey, B., Wan, K. H., Whitelaw, K. R., Yee, A., Zhang, R., Zieran, L. L.
and Kimmel, B.
Direct Submission
Submitted (20-JUN-1998) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site
(http://fruitfly.berkeley.edu/sequence-archive.html) or send email
to drosophila@mcg.lbl.gov.
Library locations: 4_36, 31_31.
Location/Qualifiers
1. 98942
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
/map="24A1-24C5"
/clone="P1s DS03364 (D237) extends from bp 1 to bp 79652 and
/note="DS03364 (D237) extends from bp 15716 to bp 98943"
DS02511 (D145) extends from bp 20166 g 28603 t
BASE COUNT 29273 a 20900 c 20166 g 28603 t
ORIGIN

alignment_scores:
Quality: 131.00 Length: 56
Ratio: 3.639 Gaps: 0
Percent Similarity: 64.286 Percent Identity: 41.071

alignment_block:
US-09-215-435-177_COPY_25_99 x AC005149/rev ..
Align seg 1/1 to reverse of: AC005149 from: 1 to: 98942
15 AspProCysLysLeuAspMetAsnPhcGlySerCysTyrGluValHisPsh 31
::: ||| :::: ||| ||::: :::::
2464 GAGGACTGTCTATCACCACCAAGAACCGGTGCTGTTTCGCTCTGTTCTA 2415
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
::: ||| ||::: ::::: ||| ||::: ||| ||:::
2414 TCGCTACGCTACACGCTGATACACAACTCTCGAGGAGTTCGTTTACG 2365
48 erGlyCysAsnGlyAsnLeuAsnAsnPhcLysLeuLysLileGluArgGlu 64
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2364 GTGGATGTCCCGCAACAGAACAACTTCGATCCAGGAGCAGCAGTGCAG 2315
65 ValAlaCysValAlaLys 70
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2314 CAGCATGTTTGTGTAAG 2297

seq_name: gb_inl:AE003579
seq_documentation_block:
LOCUS AE003579 304383 bp DNA INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386046 section 12
of 16, complete sequence.
ACCESSION AE003579 AE002638
VERSION AE003579.2 GI:10727349
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 304383)
AUTHORS Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D.,

```



```
VSTAFLEGHGIGSTLLENQVLLPQNAFIQRLVDTTCSNQPMILSYCKTLLAILWGGPE
IGNLNTLPOIATAHTPAGVSSNOAIHVIOQYASNDRLYDWMGSKRNLLEYGVSEPPA
YDLKNTITSELYLYGLADGSSANKODISRLPDLNLLHEVPDSTWGHLDIFIATEV
KRVINDVLVDISKAYDARETNEI"
join(18041..18296,18504..18607,18677..19562,19620..19772,
19836..199886)
/sequence="CG2774"
/feature="CG2774"
/feature="CT9125"
/db_xref="FLYBASE:FBan0002774"
/db_xref="FLYBASE:FBgn0031534"
<18041..>199886
/gene="CG2774"
/map="23F1-23F3"
/db_xref="FLYBASE:FBan0002774"
/db_xref="FLYBASE:FBgn0031534"
join(18114..18296,18504..18607,18677..19562,19620..19772,
19836..199886)
/sequence="CG2774"
/feature="CG2774"
/feature="gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0002774"
/db_xref="FLYBASE:FBgn0031534"
/protein_id="AAF51109.1"
/db_xref="GI:7295807"
/translation="MEVESPEHTRDFAEVDINNGAATGSEDEEVEVPAGSVTLDRNE
SDLFVSALSPSSIGDIHPLQEVLTDDGYFISIVVSDPOKIGDMGSLYAKVTTKTN
IPKFRSFSLRFSDFGLHDLVLGKYMPLGRILIPAPSKNIIGSKVKISQOSE
PCTPWTQEWIRRAALERFVHRTAQHPVLVLDLDFMFLSDQLPRSVNTSALSGA
VILRLFNKVGTVNTIKMDENPDWFDKITEVESLDANLQKLNAMKSLVTSRREL
SLTLFLNKAAMSLCTDEHTGLSRALSNLADVEEKIELLRSEANSDFFTLAEFKD
YLGFLGAIKCFHERVKAFOQWYQAOLSKRRENRRGRFELANRADKLDDQAQOEVDW
QKQVRCQQOQDDISAEIKREMERFELTRVDFKVNLIKIEDQMAHQOQIVSYWEAF
APFAREIV"
join(20807..20990,21059..21597,21748..>21915)
/sequence="CG12795"
/feature="CG12795"
/feature="CT38238"
/db_xref="FLYBASE:FBan0012795"
/db_xref="FLYBASE:FBgn0031535"
<20807..>21915
/gene="CG12795"
/map="23F1-23F3"
/db_xref="FLYBASE:FBan0012795"
/db_xref="FLYBASE:FBgn0031535"
join(20936..20990,21059..21597,21748..21915)
/sequence="CG12795"
/feature="CG12795"
/feature="gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0012795"
/db_xref="FLYBASE:FBgn0031535"
/protein_id="AAF51108.1"
/translation="MEPFLGHQHCSEATCNRLDPLPVKCDSCDKVFCASHSYDRHSC
PGATKKNQVPCVCLRPVPTPPGVEPDVTYQGHIDQCKSESKKIYTNRCNAGCK
RELIPVSCQCNFLCRLHRTSDHCKPVPASTSTSSRGGFOSIFKTSDDSRMAA
QAAERQNRFPASISNTNSRPVPQATQVQNIQGNSEDEALARALALSTIMEQDDA
AESNRQAPAPNAQOAVGVGGNOQNGKDKCLLS"
complement(join(<22424..22619,22678..23381,23475..23604,
23663..23823,23883..24022,24076..24563,24621..25106,
25190..>25354))
/sequence="CG3248"
/feature="CG3248"
/feature="CT10923"
/db_xref="FLYBASE:FBan0003248"
/db_xref="FLYBASE:FBgn0031536"
complement(<22242..>25354)
/gene="CG3248"
/map="23F1-23F3"
/db_xref="FLYBASE:FBan0003248"
/db_xref="FLYBASE:FBgn0031536"
complement(join(22242..22619,22678..23381,23475..23604,
23663..23823,23883..24022,24076..24563,24621..25106,
25190..>25354))
/sequence="CG3248"
/feature="CG3248"
/feature="gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003248"
/db_xref="FLYBASE:FBgn0031536"
/translation="MDDVERIQSENEINLRKIRNRLMOWESKTDPLAALSIQOEHLVD
LTNLWRDPSVSGPTDSAAATSQSGDDIRLPAEGCLQNTNFBLLWFADVSAIEQORG
ADYHKYLOOLQORKAECSHMLDQIAGAMERLGCALCDEYDFYSQKTSALNTSEQLIEE
QERLOELSEIQRRLHVPFVSELLNRLQSPSLVSAFARECLNKIDECNLNTEPNP
KFDAAAYNVKXROCLAKASGLVNRVTSYVINAQTEATLHPKNNPDSAAALKAPDAA
FALYKGYOTAAKRVQAQIESRHSLSYLAQMLQQHYLAQRASVSPAVNLS
TONPKVAHKGDHCSLTSRACAFVHVCODEQRLFYQFFSTGAPHTLVYLEGCLTLYD
TRMPFIIHINLETIAEISLRIEMLEHHVQNVQVALEAFATTAHOLLQDVOERLVD
RAHLYLQSDIQNFNSGDLAYPEKLEMESIALSLOPAPLRRRSDSNVSVSSA
VETESVATYTKOMIRRTVCLSLRFLYCDVRPIFOGLSOBALKICIOSVSHAACKIS
ANKTPIDGELFEIKHLLILREQIAPFRVDFVTKETSLDFSKVKTAAGFLLQKRLQFS
MGSNNALLEFLEETPQIKHLLDSRKEVDRLQKLSVCEIKIDAVOMLVGLPTFLFK
AQSLLAQSTPATQSPSTKASYVLQRWSPASPOQISIIQETORLAKIAKLAVLQSRM
QYLSNRDTEFIIFRIRNNITQSFVKLEQLTTNGYSTDDMIITSCPSABQVSILLS
SASILAEGVASFAAAARKISTSSVEGSTVGRKLSAMSNKSELVEEPKVEEVAQIE
VVAEVOPPAKIEEEVETOTQANSE"
join(25648..25833,25904..26518,26573..28394,28452..28590)
/sequence="CG8843"
/feature="CG8843"
/feature="CT9189"
/db_xref="FLYBASE:FBan0008843"
/db_xref="FLYBASE:FBgn0031537"
<25648..>28603
/gene="CG8843"
/map="23F1-23F3"
/db_xref="FLYBASE:FBan0008843"
/db_xref="FLYBASE:FBgn0031537"
join(25725..25833,25904..26518,26573..28394,28452..28590)
/sequence="CG8843"
/feature="CG8843"
/feature="gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0008843"
/db_xref="FLYBASE:FBgn0031537"
/protein_id="AAF51106.1"
/db_xref="GI:7295804"
/translation="MAPQPVTVGLSPKPGPPGTRVIRGEFLGTRVQDLGLKICGSD
CLLSAEMKSPNKIARTGPAKKGDIIVTTLSSGGVGTSTVQFRATHETIGPLKESAVV
alignment_scores:
Quality: 131.00 Length: 56
Ratio: 3.639 Gaps: 0
Percent Similarity: 64.286 Percent Identity: 41.071
alignment_block:
US-09-215-435-177_COPY_25_99 x AE003579/rev ..
Align seg 1/1 to reverse of: AE003579 from: 1 to: 304383
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
::: ||| ::| ||| |||::: ::|
265613 GAGGACTGTTCATCAGCCCAAGAAACCGGTGCGTGTTCGCTCTGTTCTA 265564
31 eATgTyRPhETyRAsnArghThrSerLysArGcysGluThrPheVaIPheS 48
::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265563 TCCTACGCTACCAACCGTGGATACACATCTCGGAGGAGTTCGTTTACG 265514
48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265513 GTGGATGTGCGGCAACACAGCAACTTCGATCCAAAGGAGCAGTGCAGAA 265464
65 ValalaCysValalaLys 70
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265463 CAGGCATGTTTGTGTAAG 265446
seq_name: gb_om:MEELACP
seq_documentation_block:
LOCUS MEELACP 419 bp mRNA MAM 15-APR-1998
DEFINITION Macropus eugenii mRNA for early lactation protein (ELP).
```

167 GCAATGGCAATAGGAACATTTT 189

seq_name: gb_pat2:163559

seq_documentation_block: 1455 bp DNA PAT 07-OCT-1997

LOCUS I63559 Sequence 6 from patent US 5663143.

DEFINITION I63559

ACCESSION I63559

VERSION I63559.1 GI:2481132

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1455)

AUTHORS Ley,A.Charles, Ladner, R.Charles, Guterman,S.Kosov, Roberts,B.Lindsay, Markland,W and Kent,R.Baribault.

TITLE Engineered human-derived Kunitz domains that inhibit human neutrophil elastase

JOURNAL Patent: US 5663143-A 6 02-SEP-1997;

FEATURES

source 1..1455

BASE COUNT 337 a 316 c 366 g 436 t

ORIGIN

alignment_scores:

Quality: 130.00 Length: 72

Ratio: 3.333 Gaps: 2

Percent Similarity: 54.167 Percent Identity: 41.667

alignment_block:

US-09-215-435-177_COPY_25_99 x I63559 ..

Align seg 1/1 to: I63559 from: 1 to: 1455

11 GlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCystY 27

49 GCGGCCAAAGAGACTCTTGCACGCTACGCGCCGCTCCCTGCAT 98

27 rgluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysGluT 44

99 GGGATGACCACAGGTATTTCTATATATGTTACATCCATGCGCTGTGAGA 148

44 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLys 60

149 CTTTCAGTACGCGCGCTGATGCGCAACGGTACACACTTCGTACACAGAA 198

61 IleGlu.....ArgGluVal..... 65

199 AGGAGTGTCTGCAGACCTGCGCAACTGTGGCGCGCTGAAACTGTTTGA 248

66 .AlaCysValAlaLys 70

249 AAGTTGTTTAGCAAAA 264

seq_name: gb_pr9:HSAIMBG3

seq_documentation_block: 4627 bp DNA PRI 06-OCT-1999

LOCUS HSAIMBG3

DEFINITION Human gene for alpha-1-microglobulin-bikunin, exons 7-10 (encoding linker peptide and bikunin).

ACCESSION X54818

VERSION X54818.1 GI:28575

KEYWORDS alpha-1-microglobulin; bikunin; glycoprotein; proteinase inhibitor.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4627)

AUTHORS Vettr,H. and Gebhard,W.

TITLE Structure of the human alpha 1-microglobulin-bikunin gene

JOURNAL Biol. Chem. Hoppe-Seyler 371 (12), 1185-1196 (1990)

AJ000490

VERSION AJ000490.1 GI:3059123

KEYWORDS early lactation protein; ELP gene.

SOURCE tammar wallaby.

ORGANISM Macropus eugenii

REFERENCE 1 (bases 1 to 419)

AUTHORS Simpson,K.J.

JOURNAL Direct Submission

TITLE Submitted (17-JUL-1997) Simpson K.J., Division of Molecular Biology and Genetics, Victorian Institute of Animal Science, 475 Mickieham Rd, Attwood, Victoria, 3049, AUSTRALIA

REMARK Revised by [3]

REFERENCE 2 (bases 1 to 419)

AUTHORS Simpson,K.J., Shaw,D. and Nicholas,K.R.

TITLE Asynchronous expression of a putative protease inhibitor gene in the mammary gland during lactation in the Tammar wallaby, Macropus eugenii

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 419)

AUTHORS Simpson,K.J.

JOURNAL Direct Submission

TITLE Submitted (14-APR-1998) Simpson K.J., Division of Molecular Biology and Genetics, Victorian Institute of Animal Science, 475 Mickieham Rd, Attwood, Victoria, 3049, AUSTRALIA

COMMENT On Apr 18, 1998 this sequence version replaced gi:3046757.

FEATURES

source 1..419

/organism="Macropus eugenii"

/db_xref="taxon:9315"

1..252

/gene="ELP"

<1..252

/gene="ELP"

/codon_start=1

/evidence="experimental

/product="early lactation protein"

/protein_id="CAA04128.1"

/db_xref="GI:3046758"

/db_xref="SPTREMBL:O62845"

/translation="EKLDQIQVNSLESLVSLCLLPVVRGNCSSQILHYFYNTT SRCCEFLVSGCNGNRRNFSEYCLKTRKNRNNNN"

1..255

/organism="Macropus eugenii"

/db_xref="taxon:9315"

/tissue_type="lactating mammary gland"

370..375

polyA_signal 139 a 98 c 72 g 110 t

polyA_site 386

BASE COUNT 139 a 98 c 72 g 110 t

ORIGIN

alignment_scores:

Quality: 130.00 Length: 41

Ratio: 4.333 Gaps: 0

Percent Similarity: 73.171 Percent Identity: 56.098

alignment_block:

US-09-215-435-177_COPY_25_99 x MEELACP ..

Align seg 1/1 to: MEELACP from: 1 to: 419

17 CysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyr 33

67 TGTCTGCTCCCACTGTGAGGGGTACTGCAGTCTCAATTTCTACATA 116

33 rPheTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyC 50

117 TTTTCAACACACACCTCTCGAACCTGTGAGACTTTCATCTACAGTGCT 166

50 ysAsnGlyAsnLeuAsnAsnPhe 57

```

MEDLINE          91214554
REFERENCE        2 (bases 1 to 4627)
AUTHORS          Gebhard, W.
TITLE            Direct Submission
JOURNAL          Submitted (09-OCT-1990) Gebhard W., Universitaet Muenchen, Klinikum
                  Grosshadern, HNO Forschungslabor, Marchioninistr 15, 8000 Muenchen
                  70. Germany
COMMENT          Sequence is approximately 6 kbp downstream of x54817.
FEATURES         Location/Qualifiers
                  1..4627
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /cell_type="leukocyte"
                    /clone_lib="EMBL3 vector"
prlm_transcript  <1..2801
                  /gene="alpha-1-microglobulin-bikunin"
                  <1..171
intron           /gene="alpha-1-microglobulin-bikunin"
                  /number=6
gene             1..2801
                  /gene="alpha-1-microglobulin-bikunin"
exon             172..253
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=7
                  /product="linker peptide and bikunin"
                  join(172..253,1339..1506,1832..2005,2658..2801)
                  /partial
mRNA             /gene="alpha-1-microglobulin-bikunin"
                  /usedin=X54816:A1MB_mRNA_a
                  /usedin=X54816:A1MB_mRNA_b
                  join(172..253,1339..1506,1832..2005,2658..2801)
                  /partial
                  /gene="alpha-1-microglobulin-bikunin"
                  /note="alpha-1-microglobulin"
                  /codon_start=1
                  /usedin=X54816:A1MB_cds
                  /product="bikunin"
                  /protein_id="CAA38587.1"
                  /db_xref="GI:579676"
                  /db_xref="SPTREMBL:P78491"
                  /translation="RVRRAVLQPEEGSGGQLYTEVTKKEDSQLGYSAGPCMGMTS
                  RYFYNGTSMACETQYGGCMGNNFVTEKELQCTVAAACNLPVIRGPCRAFTQLW
                  AFDVAKGVCLFFPYGGCGNGKFYSEKREYCGVPGDGEELLRFNS"
                  254..1338
intron           /gene="alpha-1-microglobulin-bikunin"
                  /number=7
exon             1339..1506
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=8
intron           1507..1831
                  /product="bikunin"
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=8
exon             1832..2005
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=9
                  /product="bikunin"
intron           2006..2657
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=9
exon             2658..2801
                  /gene="alpha-1-microglobulin-bikunin"
                  /number=10
                  /product="bikunin"
polyA_signal     2773..2778
                  /gene="alpha-1-microglobulin-bikunin"
                  2801
polyA_site       /gene="alpha-1-microglobulin-bikunin"
                  4288..4612
repeat_unit      /rpt_family="Alu"
BASE COUNT      1088 a 1133 c 1292 g 1108 t 6 others
ORIGIN

```

```

alignment_scores:
  Quality: 129.50          Length: 59
  Ratio: 3.597            Gaps: 1
  Percent Similarity: 61.017 Percent Identity: 45.763

alignment_block:
US-09-215-435-177_COPY_25_99 x HSAIMBG3 ..
Align seg 1/1 to: HSAIMBG3 from: 1 to: 4627

9 ILeCysGlyAspLeuTysAspProCysLysLeuAspMetAsnPhGlySe 25
   :::::  :::::  |||  |||  :::::  :::::  :::::  |||
1323 GTCTGCCCTCCCTCA...GATTCCTCCAGCTGGGCTACTCGGCCGCTCC 1369

25 rCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgC 42
   |||  :::::  |||  |||  |||  |||  |||  |||  |||  |||
1370 CTCATGGGAATGACCAGCAGGATTTCTATAATGGTACATCCATGCCCT 1419

42 ySGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhLys 58
   |||  |||  |||  |||  :::::  |||  |||  |||  |||  |||
1420 GTGAGACTTTCAGTACGGCGGCTGCATGGCAACGGTAACAACACTTCGTC 1469

59 LeuLysileGluArgGluValAlaCys 67
   :::::  |||  :::::
1470 ACAGAAAGGAGTGTCGTCAGACCTGC 1496

seq_name: gb_pr10:HUMITILC08

seq_documentation_block:
LOCUS      HUMITILC08      5423 bp      DNA      PRI      06-JAN-1995
DEFINITION Human inter-alpha-trypsin inhibitor light chain (IT1) gene, exons
            7-10.
ACCESSION  M88249.1 GI:186598
VERSION    M88249.1
KEYWORDS   inter-alpha-trypsin inhibitor light chain.
SEGMENT    8 of 8
SOURCE     Homo sapiens (tissue library: of T.Maniatis and Clontech) fetus
            liver DNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 5423)
AUTHORS    Diarra-Mehrpour, M., Bourguignon, J., Sesboue, R., Salier, J.P.,
            Leveillard, T. and Martin, J.P.
TITLE      Structural analysis of the human inter-alpha-trypsin inhibitor
            light-chain gene
JOURNAL    Eur. J. Biochem. 191 (1), 131-139 (1990)
MEDLINE    E09336621
FEATURES   Location/Qualifiers
            1..5423
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /dev_stage="fetus"
              /tissue_type="liver"
              /tissue_lib="of T.Maniatis and Clontech"
              /map="9q32-q33"
              join(M88165.1:532..1176,M88243.1:1..489,M88244.1:1..618,
              M88245.1:1..340,M88246.1:1..690,M88247.1:1..335,1..3604)
              /gene="IT1 light chain"
              join(M88165.1:690..900,M88243.1:265..407,
              M88244.1:323..399,M88245.1:198..314,M88246.1:138..239,
              M88247.1:192..238,976..1057,2145..2312,2638..2811,
              3464..3604)
              /gene="IT1 light chain"
              /product="inter-alpha-trypsin inhibitor light chain"
              join(M88165.1:690..1176,M88243.1:1..489,M88244.1:1..618,
              M88245.1:1..340,M88246.1:1..690,M88247.1:1..335,
              M88248.1:1..287,1..3495)
              /gene="ITIL"
              join(M88165.1:694..900,M88243.1:265..407,
              M88244.1:323..399,M88245.1:198..314,M88246.1:138..239,
              M88247.1:192..238,976..1057,2145..2312,2638..2811,

```

```

0      /gene="ITI light chain"/
0      /product="inter-alpha-trypsin inhibitor light chain"
0      Join(M88165.1:784..900,M88243.1:265..407,.239,
0      M88244.1:323..399,M88245.1:198..314,M88246.1:138..239,
0      M88247.1:192..238,976..1057,2145..2312,2638..2811,
0      3464..3495)
0      /gene="ITIL"
0      /codon_start=1
0      /db_xref="GDB:M00-120-696"
0      /product="inter-alpha-trypsin inhibitor light chain"
0      /protein_id="AAA59136.1"
0      /db_xref="GI:186600"
0      translation="MRSIGALLLLLSACLAVSGPVPTPDNIQVOENNISRIYKWD
0      YNLAIQSCTPKLKMIDMTVSTLVLEGETAETASIMTSRWRKGVCSEETSGAYKETD
0      TDGFLEYHKSXNWTMSISYVVHTNYDEYAIFLTKKFSRHGGPITAKLYGRAPQLRET
0      LLQDFEVTVAQGSGVPEDSIVFTMDRGCVPEGEPEPIIPRVRRVALPOEEEGSGG
0      QLQFTEVTKKDSCLQGSAGPCMGNTSRFYNGTSMACETFXYGCCMGNGNNFVTEKE
0      CLQCRTVAACNLPIVRGPCHAFIQLWADFVGKCVLPFPYGCCQGNKFYESEKBCE
0      EYCGPGDGDEELURFS"
0      order(M88247.1:239..335,M88248.1:1..287,1..975)
0      /gene="ITI light chain"
0      /note="intron is about 5.8 kb."
0      /number=6
0      TATA_signal
0      900..906
0      /gene="ITI light chain"
0      976..1057
0      /gene="ITI light chain"
0      /number=7
0      1058..2144
0      /gene="ITI light chain"
0      /number=7
0      2145..2312
0      /gene="ITI light chain"
0      /number=8
0      2313..2637
0      /gene="ITI light chain"
0      /number=8
0      2638..2811
0      /gene="ITI light chain"
0      /number=9
0      2812..3463
0      /gene="ITI light chain"
0      /number=9
0      3464..3604
0      /gene="ITI light chain"
0      /number=10
0      repeat_unit complement(5112..5402)
0      /rpt_family="Alu repetitive element"
BASE COUNT   1258 a   1333 c   1450 g   1382 t
ORIGIN
alignment_scores:
    Quality: 129.50          Length: 59
    Ratio: 3.597            Gaps: 1
Percent Similarity: 61.017 Percent Identity: 45.763

alignment_block:
US-09-215-435-177_copy_25_99 x HUMITILC08 ..
Align seg 1/1 to: HUMITILC08 from: 1 to: 5423
9 ileCysgLyAspLeuLyAspProCysLyLysLeuAspMetAsnPheGlySe 25
::| | | | | ::| | | | | ||| | | | | | : : | | |
2129 GTCTGCCCTCCCTA...GATTCTGCACGTGGGCTACTGGCCGGTC 2175
|||||
25 rCyStyTrGLuValHisPheargTrYPheTyTAsnArgThrSerLySArgC 42
||| ::| | | | | | | | | | | | | | | | | | | | | | | | | |
2176 CTGCAATGGGAATGACCAGCACGTATTTCCTATATGGTGATCATCGCCT 2225
|||||
42 ySGluThrPheValPHeSerGlyCysAsnGlyAsnLeuAsnPhelys 58
., | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

1

```

/note="assembly_fragment"
82947..91987
/misc_feature /note="assembly_fragment"
92088..98742
/misc_feature /note="assembly_fragment"
98843..105574
/misc_feature /note="assembly_fragment"
105675..114985
/misc_feature /note="assembly_fragment"
115086..124472
/misc_feature /note="assembly_fragment"
124573..134778
/misc_feature /note="assembly_fragment"
134879..150202
/misc_feature /note="assembly_fragment"
150303..166267

alignment_scores:      Length: 59
                       Quality: 129.50
                       Ratio: 3.597
Percent Similarity:    Gaps: 1
                      Percent Identity: 45.763

alignment_block:
US-09-215-435-177_COPY_25_99 x AC011161/rev ..

Align seg 1/1 to reverse of: AC011161 from: 1 to: 200532

          9   IleCysGlyAspLeuLysAspProCysLysLeuAsnMetAsnPhleGlySe 25
             :::||| |::||| |::||| |::||| |::||| |::|||
192488 GTCTGCCCTCCCTA...GATTCTGCAGCTGGGTACTCGGCGGGTCC 192442

          25 rCysTyrrGlunValHisPheArgtyrPheTyrrAsnArghThrSerLysAsrGc 42
             ||| ::| |::||| |::||| |::||| |::||| |::|||
192441 CTGCATGGGAATGACCACGACGGTAGTTTCATAATATGGTACATCCATGGCCT 192392

          42 ysGluIrrPheValPheSerdGlyCysAsnGIynLeuAsnAsnPheLys 58
             ||||| ||||| |::||| |::||| |::||| |::||| |::|||
192391 GTGNAGCTTTCCAGTACCGCGGCTGCATGGCGAACGGTAACAACCTTCGTC 192342

          59 LeuLysllelgluArgGluvalAlacys 67
             ::| ||| |::||| |::||| |::||| |::||| |::|||
192341 ACAGAAAAGGAGTGCTGTCAGACACTGCG 192315
```

OM of: US-09-215-435-177_COPY_25_99 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:28 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-CAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY_25_99

Query length: 75

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97612		419.00	997.35	1.4e-47	526
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97738		409.00	972.94	3.2e-46	539
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ00455		307.00	733.42	7.0e-33	267
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ15699		307.00	733.28	7.1e-33	271
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ32013		133.00	262.21	1.2e-06	42521
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ90070		133.00	262.21	1.2e-06	42521
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ35164		130.00	304.97	5.1e-09	285
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAV40046		128.50	296.98	1.4e-08	441
/cgnl_9/gcgdata/geneseq/geneseq/NA1988.DAT:AAQ81432		128.50	286.72	5.3e-08	1232
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ78950		128.50	286.53	5.5e-08	1255
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ77939		128.50	285.13	6.5e-08	1444
/cgnl_9/gcgdata/geneseq/geneseq/NA1990.DAT:AAQ06878		127.00	301.71	7.8e-09	191
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46936		127.00	293.89	2.1e-08	418
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79081		126.00	292.64	2.5e-08	372
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79088		126.00	290.37	3.3e-08	467
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ15870		125.00	295.38	1.8e-08	222
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79079		124.00	295.93	1.6e-08	165
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79080		124.00	287.80	4.6e-08	372
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79083		124.00	286.24	5.7e-08	435
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79086		124.00	285.53	6.2e-08	467
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79087		124.00	284.27	7.3e-08	530
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46935		123.00	284.22	7.3e-08	418
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79077		122.00	291.85	2.8e-08	153
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ86987		121.00	275.89	2.1e-07	593
/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ90469		121.00	270.89	4.1e-07	979
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46934		120.00	276.97	1.9e-07	418
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ51001		120.00	268.77	5.3e-07	950
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46933		119.00	274.55	2.5e-07	418
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ29017		118.00	280.55	1.2e-07	180
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ29024		118.00	277.19	1.8e-07	252
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ33175		118.00	274.03	2.7e-07	346
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ32238		118.00	262.34	1.2e-06	1114
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ35169		117.00	272.90	2.4e-07	249
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ10945		117.00	261.99	1.9e-06	906
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ39947		117.00	258.95	1.3e-06	1228
/cgnl_9/gcgdata/geneseq/geneseq/NA1991.DAT:AAQ10992		116.00	266.05	7.6e-07	474
/cgnl_9/gcgdata/geneseq/geneseq/NA1991.DAT:AAQ10993		116.00	262.25	1.2e-06	693
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ51316		116.00	259.51	1.7e-06	912
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ51317		116.00	259.51	1.7e-06	912
/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ81396		116.00	259.33	1.8e-06	928

/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46644 + 116.00 259.15 1.8e-06 945
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ46931 + 116.00 259.15 1.8e-06 945
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ16056 + 116.00 257.96 2.1e-06 106
/cgnl_9/gcgdata/geneseq/geneseq/NA1989.DAT:AAQ90108 + 116.00 255.01 3.1e-06 143
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ32167 + 116.00 255.01 3.1e-06 143

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ97612

seq_documentation_block:

ID_AAX97612 standard; DNA; 526 BP.

AC AAX97612;

DT 13-SEP-1999 (first entry)

XX Extended human secreted protein coding sequence, SEQ ID NO. 76.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.

XX Homo sapiens.

OS Homo sapiens.

PN WO9931236-A2.

XX 24-JUN-1999.

PD 17-DEC-1998; 98WO-IB02122.

XX 10-AUG-1998; 98US-00961116.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0061563.

XX (GEST) GENSET.

PI Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI; 1999-385906/32.

DR P-PSDB; AAY35928.

XX New isolated human secreted proteins

PS Claim 1; Page 210; 516pp; English.

XX This sequence represents an extended human secreted protein coding
sequence of the invention. The secreted proteins can be used in treating
or controlling a variety of human conditions. The secreted proteins may
act as cytokines or may affect cellular proliferation or differentiation
or may act as immune system regulators, haematopoiesis regulators, tissue
growth regulators, regulators of reproductive hormones or cell movement
or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
tumour inhibition activity. The DNAs can be used in forensic procedures
to identify individuals or in diagnostic procedures to identify
individuals having genetic diseases resulting from abnormal expression of
the genes corresponding to the extended cDNAs. They are also useful for
constructing a high resolution map of the human chromosomes. They can
also be used for gene therapy to control or treat genetic diseases.

SQ Sequence 526 BP; 162 A; 100 C; 99 G; 165 T; 0 other;

alignment_scores:

Quality: 419.00 Length: 75
Ratio: 5.587 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-177_COPY_25_99 x AAX97612

Align seg 1/1 to: AAX97612 from: 1 to: 526

Thu Nov 15 10:51:57 2001

CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.

XX
SQ sequence 539 BP; 165 A; 103 C; 97 G; 162 T; 12 other;

alignment_scores:
Quality: 409.00 Length: 75
Ratio: 5.453 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.333

alignment_block:

US-09-215-435-177_COPY_25_99 x AAX97738 ..

Align seg 1/1 to: AAX97738 from: 1 to: 539

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
|||||
104 GGTGTTAATAAAATTCGGAGAAGATATGTCGAGACCTCAAGATCCCTG 153
17 sLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrP 34
|||||
154 CAAATTGGACATGAATTTTGGAGCTGCTATGAAGTTCACTTTAGATATT 203
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
|||||
204 TCTACACAGAACCTCCAAAAGATGTGAAACTTTTGTCTTCCAGCTGT 253
51 AsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCys 67
|||||
254 AATGGCAACCTTAACAACTTCAAGCTTAATAATAGAACGTGAAGTAKCCTG 303

67 sValAlaLysTyrLysProProArg 75

304 TGTTCGCAAAATACAAACCCAGG 328

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAC00455

seq_documentation_block:

ID AAC00455 standard; cDNA; 267 BP.

XX AAC00455;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 453.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX P-PSDB; AAC00449.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures
XX Claim 1; SEQ ID 453; 71pp + CD-ROM; English.

XX

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
|||||
94 GGTGTTAATAAAATTCGGAGAAGATATGTCGAGACCTCAAGATCCCTG 143
17 sLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrP 34
|||||
144 CAAATTGGACATGAATTTTGGAGCTGCTATGAAGTTCACTTTAGATATT 193
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
|||||
194 TCTACACAGAACCTCCAAAAGATGTGAAACTTTTGTCTTCCGGCTGT 243
51 AsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCys 67
|||||
244 AATGGCAACCTTAACAACTTCAAGCTTAATAATAGAACGTGAAGTACCTG 293
67 sValAlaLysTyrLysProProArg 75
|||||
294 TGTTCGCAAAATACAAACCCAGG 318

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAX97738

seq_documentation_block:

ID AAX97738 standard; DNA; 539 BP.

XX AAX97738;

XX 13-SEP-1999 (first entry)

XX Extended human secreted protein coding sequence, SEQ ID NO. 303.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
XX cellular differentiation; immune system regulator; anti-inflammatory;
XX haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
XX reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
XX genetic disease; ss.

XX Homo sapiens.

XX WO9931236-A2.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-IB02122.

XX 10-AUG-1998; 98US-0096116.

XX 17-DEC-1997; 97US-0069957.

XX 09-FEB-1998; 98US-0074121.

XX 13-APR-1998; 98US-0081563.

XX (GEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI: 1999-385906/32.

XX P-PSDB; AAY36054.

XX New isolated human secreted proteins

XX Claim 1; Page 376; 516pp; English.

XX This sequence represents an extended human secreted protein coding
XX sequence of the invention. The secreted proteins can be used in treating
XX or controlling a variety of human conditions. The secreted proteins may
XX act as cytokines or may affect cellular proliferation or differentiation
XX or may act as immune system regulators, haematopoiesis regulators, tissue
XX growth regulators, regulators of reproductive hormones or cell movement
XX or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
XX tumour inhibition activity. The DNAs can be used in forensic procedures
XX to identify individuals or in diagnostic procedures to identify
XX individuals having genetic diseases resulting from abnormal expression of
XX the genes corresponding to the extended cDNAs. They are also useful for

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

XX Sequence 267 BP; 74 A; 57 C; 52 G; 84 T; 0 other;

alignment_scores:
 Quality: 307.00 Length: 55
 Ratio: 5.582 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.182

alignment_block:

US-09-215-435-177_COPY_25_99 x AAC00455 ..

Align seg 1/1 to: AAC00455 from: 1 to: 267

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
 |||||
 103 GGTGTTAATAAAATTCGGGAGAGATATGTCGAGACCTCAAAGATCCCTG 152
 17 sLysLeuAspMetAsnPhcGlySerCysTyrGluValHisPheArgTyrP 34
 |||||
 153 CAAATGGACATGAATTTGGAGAGCTGCTATGAGTTCACTTTAGATATT 202
 34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
 |||||
 203 TCTACACACAGAACCTCCAAAAGATGTGAAACTTTTGTCTCTCCAGCTGT 252
 51 AsnGlyAsnLeuAsn 55
 |||||
 253 AATGGCAACCTTAAC 267

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAx51569

seq_documentation_block:

ID AAX51569 standard; cDNA; 271 BP.

XX AC AAX51569;

DT 21-JUN-1999 (first entry)

XX DE Human secreted protein 5' EST SEQ ID NO:148.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX OS Homo sapiens.

XX XX WO9906549-A2.

XX XX 11-FEB-1999.

XX XX 31-JUL-1998; 98WO-IB01231.

XX XX 01-AUG-1997; 97US-0905279.

XX XX (GEST) GENSET.

XX

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

DR WPI: 1999-153779/13.

DR P-PSDB; AAY12791.

XX

PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from testis, ovary, uterus and spleen tissue
 XX
 PS Claim 1; Page 253-254; 522pp; English.

XX

CC AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12681 to
 CC AAY12913, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

SQ Sequence 271 BP; 75 A; 57 C; 55 G; 84 T; 0 other;

alignment_scores:

Quality: 307.00 Length: 55
 Ratio: 5.582 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.182

alignment_block:

US-09-215-435-177_COPY_25_99 x AAX51569 ..

Align seg 1/1 to: AAX51569 from: 1 to: 271

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
 |||||
 104 GGTGTTAATAAAATTCGGGAGAGATATGTCGAGACCTCAAAGATCCCTG 153
 17 sLysLeuAspMetAsnPhcGlySerCysTyrGluValHisPheArgTyrP 34
 |||||
 154 CAAATGGACATGAATTTGGAGAGCTGCTATGAGTTCACTTTAGATATT 203
 34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
 |||||
 204 TCTACACACAGAACCTCCAAAAGATGTGAAACTTTTGTCTCTCCAGCTGT 253
 51 AsnGlyAsnLeuAsn 55
 |||||
 254 AATGGCAACCTTAAC 268

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ32013

seq_documentation_block:

ID AAZ32013 standard; DNA; 42521 BP.

XX AC AAZ32013;

XX DT 10-JAN-2000 (first entry)

XX DE Human METH1 related EST U64857.

XX Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
 KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
 KW angiogenesis inhibitor; abnormal wound healing; inflammation;
 KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;

CC fibromuscular dysplasia, wound granulation, Crohn's disease or
 CC atherosclerosis. METH can also be used in birth control. METH can also
 CC be used in diagnostic methods for the prognosis of cancer.
 XX
 SQ Sequence 42521 BP; 13493 A; 8055 C; 8365 G; 12608 T; 0 other;

alignment_scores:
 Quality: 133.00 Length: 62
 Ratio: 3.500 Gaps: 0
 Percent Similarity: 61.290 Percent Identity: 40.323

alignment_block:

US-09-215-435-177_COPY_25_99 x AAC90070 ..

Align seg 1/1 to: AAC90070 from: 1 to: 42521

10 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheCysSerCy 26
 ||| :: ||||| |||::||| ||| ||
 6370 TGTCAAATCATAGAGTCTTGTCAACTTCCAAGTTCAAGGACCATG 6419

26 sTyrgluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 43
 ||| |||::||| |||::||| |||::||| |||::|||

6420 CTCTGGAAGCATCTCTATTATTACTACACACTGCCAGTTCATCAATGCG 6469

43 luThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeu 59
 |||||::|||::|||::||| ||||| |||::||| |||::|||

6470 AGAGGTTCACTATGTGTGCTGCTGGGAATACTAACAGATTCGTACC 6519

60 LysIleGluArgGluValAlaCysValAlaLysTyr 71
 ||| ::||| ||| ::||| ||| ::||| ||| ::|||

6520 ATTGAGGAGTGTCAAGCAGATGCCGAGTAAGTTC 6555

seq_name: /cgn1_9/gcdata/geneseq/geneseq/NA1996.DAT:AAT35164

seq_documentation_block:

ID AAT35164 standard; DNA; 285 BP.

XX AC AAT35164;

XX 14-FEB-1997 (first entry)

DE ITI-D1 Kunitz domain fusion gene.

XX Aprotinin; Kunitz domain; human neutrophil elastase; hNE;
 KW connective tissue; alpha 1 protease inhibitor; API; neutrophil;
 KW alaphal antitrypsin; respiratory disorder; cystic fibrosis;
 KW smokers emphysema; ss.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 1..285

FT /*tag= a

FT /product= IIISp::itiD1::matureIII (initial fragment)

FT fusion protein

FT misc_feature 1..3

FT /*tag= b

FT /transl_except= GTG encodes Met

FT sig_peptide 1..54

FT /*tag= c

FT /note= "M13 III protein signal peptide"

FT mat_peptide 55..234

FT /*tag= d

FT /note= "ITI-D1 Kunitz domain"

FT mat_peptide 235..285

FT /*tag= e

FT /note= "M13 III protein"

XX W09620278-A2.

XX 04-JUL-1996.

XX

PF 15-DEC-1995; 95WO-US16349.
 XX
 PR 16-DEC-1994; 94US-0358160.
 XX
 PA (PROT-) PROTEIN ENG CORP.
 XX
 PI Guterman SK, Kent RB, Ladner RC, Ley AC, Markland W;
 PI Roberts BL;
 XX
 DR WPI: 1996-321851/32.
 DR P-FSDB; AAR99213.
 XX
 PT New engineered inhibitors of human neutrophil elastase - contg.
 PT aprotinin-like Kunitz domain for treating, e.g. cystic fibrosis or
 PT other respiratory disorders
 XX
 PS Example 1; Page 43; 105pp; English.
 XX
 CC Genetically engineered human derived Kunitz domains can be used to
 CC inhibit human neutrophil elastase, an enzyme involved in the
 CC elimination of pathogens and the restructuring of connective tissue.
 CC In cases of reduction of the circulating alpha-1-protease inhibitor
 CC (API or alaphal antitrypsin), or the inactivation of API by oxidation
 CC (smokers emphysema), extensive destruction of the lung tissue may
 CC result from uncontrolled elastolytic activity of human neutrophil
 CC elastase. Other respiratory disorders such as cystic fibrosis are
 CC thought to be caused by human neutrophil elastase release by
 CC neutrophils. The genetically engineered human derived Kunitz
 CC domains can be used to treat such respiratory disorders. See
 CC AAR99146-R99211. Fusion genes were used in the production of the
 CC Kunitz domain derivatives. This sequence is a fragment of a fusion
 CC gene that encodes the M13 III signal peptide, ITI-D1 Kunitz domain
 CC and the mature M13 III protein. Substitutions in the ITI-D1 domain
 CC can be made by use of BglI, EagI, NcoI, StylI, PstI and Kasi
 CC restriction sites found within the sequence or by
 CC oligonucleotide-directed mutagenesis of single-stranded phage DNA.
 CC The resulting peptide derivatives can be expressed as intact human
 CC neutrophil elastase binding proteins or can be incorporated into
 CC larger proteins as a domain.
 XX
 SQ Sequence 285 BP; 74 A; 67 C; 68 G; 76 T; 0 other;

alignment_scores:

Quality: 130.00 Length: 72
 Ratio: 3.333 Gaps: 2
 Percent Similarity: 54.167 Percent Identity: 41.667

alignment_block:

US-09-215-435-177_COPY_25_99 x AAT35164 ..

Align seg 1/1 to: AAT35164 from: 1 to: 285

11 GlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCysty 27

49 GCGCCCAAAGAACACTCTTCCAGCTGGGCTACTGGCGCGTCCCTGCAT 98

27 rGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysGluT 44

99 GGAATGACACGACGAGGTATTTCTATAATGTATCATCCATCGCCTGTGAGA 148

44 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLys 60

149 CTTTCAGTACGCGCGTGCATGGCAACGGTACACACTTCGTACAGAA 198

61 IleGlu.....ArgGluVal..... 65

199 AAGGAGTGTCTGCAGACCTGCCGAACCTGTGGCGCGCTGAACACTGTGTA 248

66 AlaCysValAlaLys 70

249 AAGTTGTTTAGCAAAA 264

US-09-215-435-177_COPY_25_99 x AAV40046

PT elastase p

PT elastase p

us-09-215-435-177_copy_25_99.rng

Thu Nov 15 10:51:57 2001

```

PR 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX DR P-PSDB; AAB43730.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX PT useful for treating or diagnosing e.g. cancer -
XX PS Claim 1; Page 885; 2352pp; English.
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
XX neotropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention.
XX SQ Sequence 1444 BP; 362 A; 383 C; 422 G; 277 T; 0 other;

alignment_scores:
Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x AAC77939 ..
Align seg 1/1 to: AAC77939 from: 1 to: 1444
2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLY 18
895 GTAACTGAAGTCACCAAGAAA.....GAGATTCCTGCCA 929
18 sLeuAspMetAsnPhelGlySerCysTyrGluValHisPheArgTyrPheT 35
930 GCTGGGCTACTCGCGCGTCCCTGCATGGGAATGACACGAGGTATTCT 979
35 yrAsnArgThrSerLysArgCysGlyThrPheValPheSerGlyCysAsn 51
980 ATAATGATACATCCATCGGCTGTGAGACTTTCAGTACGGCGGCTGCATG 1029
52 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 67
1030 GGCACACGGTAAACACTTCGTCACAGAAAGGAGTGCTGCAGACCTGC 1077
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA190.DAT:AAQ06878
seq_documentation_block:
ID AAQ06878 standard; DNA; 191 BP.
XX AC AAQ06878;
XX PT 05-MAR-1991 (first entry)

```

```

XX DE Encodes human bikunin domain 1-derived protease inhibitor.
XX KW bikunin; serine protease inhibitor; rheumatoid arthritis; ds.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT mutation 57..59
XX FT /*tag= a
XX FT /*note= "substitute with codon for Leu, Ile, Val,
XX FT Arg, Phe, Tyr, Thr or Lys"
XX FT mutation 63..65
XX FT /*tag= b
XX FT /*note= "substitute with codon for Leu, Arg, Ile,
XX FT Val or Lys"
XX EP401508-A.
XX 12-DEC-1990.
XX 01-MAY-1990; 90EP-0108284.
XX 18-JAN-1990; 90DE-4001244.
XX 13-MAY-1989; 89DE-3915689.
XX (FARB ) BAYER AG.
XX Fritz H, Gebhard W, Das R;
XX WPI; 1990-369615/50.
XX P-PSDB; AAR08291.
XX Protease-inhibitor variant derived from human bikunin - prodn.
XX used as therapeutic for treatment of eg. rheumatoid arthritis
XX Disclosure; Fig 2; 29pp; German..
XX The 5' end overhangs the 3' end of the complementary strand by 5
XX bases; the 5' end of the complementary strand overhangs the 3' end
XX of the sense strand by AGCTG. The sequence encodes residues 19-77 of
XX mature bikunin in which Met(36) and/or Met(38) are substituted by
XX another naturally occurring amino acid. The sequence comprises domain
XX 1 of bikunin and is synthesised from two synthetic oligonucleotides.
XX The substitutions are achieved by site-directed mutagenesis.
XX Plasmids containing the sequences are used to transform bacteria
XX and the protease inhibitory fragments are secreted as fusion
XX proteins. The fragments can be used for the treatment of emphysema,
XX septic shock, rheumatoid arthritis and coagulation disorders.
XX See also AAQ06879.
XX SQ Sequence 191 BP; 49 A; 43 C; 46 G; 53 T; 0 other;

alignment_scores:
Quality: 127.00 Length: 54
Ratio: 3.848 Gaps: 0
Percent Similarity: 61.111 Percent Identity: 46.296

alignment_block:
US-09-215-435-177_COPY_25_99 x AAQ06878 ..
Align seg 1/1 to: AAQ06878 from: 1 to: 191
14 LysAspProCysLysLeuAspMetAsnPhelGlySerCysTyrGluValHi 30
18 GAAGACTCTTGGCCAGCTGGGCTACTCTGCTGGTCCGTCATGGGTATGAC 67
30 sPheArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValP 47
68 TTCTGTTACTTCTACAAACGCTACCTCTATGCTTGGCAAACTTCAGT 117
47 heSerGlyCysAsnGlyAsnLeuAsnPhelLysLeuLysIleGluArg 63

```


us-09-215-435-177_copy_25_99.rng

Thu Nov 15 10:51:57 2001

```
US-09-215-435-177_COPY_25_99 x AAT79081 ..
Align seg 1/1 to: AAT79081 from: 1 to: 372
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
7 GATTTCTGTCAACTGGCTACTCGCGCGCTCCCTCGGTGCTATGTTTCC 56
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
57 TAGGTATTTCTATAATGGTACATCCATCGCGCTCCCTCGGTGCTATGTTTCC 106
48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
107 GCGGATGTCATGGGCAACGGTAACTTCGTCACAGAAAAGGAGTGTCTG 156
65 ValAlaCys 67
157 CAGACCTGC 165

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT.AAT79088
seq_documentation_block:
ID: AAT79088 standard; DNA; 467 BP.
XX
AC AAT79088;
XX
DT 12-NOV-1997 (first entry)
XX
DE Insert from plasmid pHH341.
XX
KW Trypsin inhibitor; Kunitz domain; protease; active site; elastase;
KW neutrophil; disease; modification; site directed mutagenesis; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 7..441
FT /*tag= a
FT /product= elastase inhibitor Ep7-d21-RPDF-52-55
FT sig_peptide 7..66
FT /*tag= b
FT mat_peptide 67..438
FT /*tag= c
XX
PN JP09124700-A.
XX
PD 13-MAY-1997.
XX
PF 07-NOV-1995; 95JP-0288527.
XX
PR 07-NOV-1995; 95JP-0288527.
XX
PA (GREC ) GREEN CROSS CORP.
XX
WPI: 1997-316576/29.
DR P-PSDB; AAW25938.
XX
PT New protease inhibitor - useful for treating diseases involving
PT elastase
XX
PS Disclosure; Fig 38; 37pp; Japanese.
XX
CC This is the nucleotide sequence of the insert for the plasmid pHH341
CC which contains the sequence encoding the novel elastase specific
CC inhibitor Ep7-d21-RPDF-52-55 (AAT79081). The inhibitor coding sequence
CC is linked downstream of the yeast invertase (SUC2) signal peptide
CC sequence. The modified protease inhibitors are targeted to the protease
CC elastase, especially from neutrophils and can be used to treat diseases
CC associated with elastase. Modifications of the active site were done by
CC site directed mutagenesis.
XX
SQ Sequence 467 BP; 90 A; 121 C; 132 G; 124 T; 0 other;
```

```
alignment_scores:
Quality: 126.00 Length: 53
Ratio: 3.818 Gaps: 0
Percent Similarity: 62.264 Percent Identity: 47.170
alignment_block:
US-09-215-435-177_COPY_25_99 x AAT79088 ..
Align seg 1/1 to: AAT79088 from: 1 to: 467
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
73 GATTTCTGTCAACTGGCTACTCGCGCGCTCCCTCGGTGCTATGTTTCC 122
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
123 TAGGTATTTCTATAATGGTACATCCATCGCGCTCCCTCGGTGCTATGTTTCC 172
48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
173 GCGGATGTCATGGGCAACGGTAACTTCGTCACAGAAAAGGAGTGTCTG 222
65 ValAlaCys 67
223 CAGACCTGC 231
```


THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-177_COPY_25_99 to: Issued_Patents_NA.* out_format : pfs

Date: Nov 15, 2001 3:58 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY_25_99

Query length: 75

Database: Issued_Patents_NA.*

Database sequences: 351203

Database length: 11323899

Search time (sec): 228.030000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-6 +	130.00	305.89	2.2e-09	1455	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-17 +	127.00	310.97	1.1e-09	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-15 +	123.00	300.66	4.3e-09	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-147-710-1 +	121.00	286.76	2.5e-08	979	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-458-090-1 +	121.00	286.76	2.5e-08	979	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-457-887-1 +	121.00	286.76	2.5e-08	979	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-817-145-2 +	121.00	285.20	3.1e-08	1140	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-13 +	120.00	292.92	1.1e-08	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-11 +	119.00	290.35	1.6e-08	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-135 +	117.00	290.51	1.6e-08	249	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-2 +	116.00	281.32	5.1e-08	474	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-4 +	116.00	277.42	8.4e-08	693	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-026-145-1 +	116.00	274.43	1.2e-07	928	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-446-846-8 +	116.00	274.43	1.2e-07	928	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-7 +	116.00	274.24	1.3e-07	945	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-854-764-1 +	116.00	273.01	1.5e-07	1065	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-09-054-782-1 +	116.00	269.98	2.2e-07	1431	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-321-658B-1 +	115.00	269.98	2.2e-07	1431	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-828-920A-6 +	114.00	287.00	2.5e-08	165	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-147-710-14 +	114.00	287.00	2.5e-08	165	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-458-090-14 +	114.00	287.00	2.5e-08	165	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-457-887-14 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-74 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-74 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-70 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-74 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-70 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-74 +	113.00	275.03	1.1e-07	412	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-084-718-58 +	111.00	269.73	2.3e-07	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-976-58 +	111.00	269.73	2.3e-07	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-443-977-58 +	111.00	269.73	2.3e-07	418	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-829-876-90 +	110.00	266.51	3.4e-07	445	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-07-985-692-1 +	107.50	261.18	6.7e-07	399	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-155-331-1 +	107.50	261.18	6.7e-07	399	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-424-022-1 +	107.50	261.18	6.7e-07	399	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-424-017B-1 +	107.50	261.18	6.7e-07	399	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-155-331-12 +	107.50	238.25	1.3e-05	3725	!
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-424-022-12 +	107.50	238.25	1.3e-05	3725	!

/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-424-017B-12 + 107.50 238.25 1.3e-05 3725
/cgnl_7/ptodata/1/ina/5B.COMB.seq:US-08-829-876-92 + 107.50 238.25 1.3e-05 3725
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-334-773A-5 + 106.00 256.84 1.2e-06 418
/cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-384-489-8 + 106.00 256.69 1.2e-06 424
seq_name: /cgnl_7/ptodata/1/ina/5A.COMB.seq:US-08-358-160-6

seq_documentation_block:
; Sequence 6, Application US/08358160
; Patent No. 5663143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/358.160
; APPLICATION NUMBER: US/08/358.160
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,031
; FILING DATE: 13-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,319
; FILING DATE: 26-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/664,989
; FILING DATE: 01-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,063
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/240,160
; FILING DATE: 02-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEV=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1455 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: synthetic DNA fragment
US-08-358-160-6
alignment_scores:
Quality: 130.00 Length: 72

```

Ratio: 3.333      Gaps: 2
Percent Similarity: 54.167      Percent Identity: 41.667

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-358-160-6      ..

Align seg 1/1 to: US-08-358-160-6 from: 1 to: 1455

11 GlyAspLeuLyAspProCysLyLeuAspMetAsnPheGlySerCysTy 27
111      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
49 GCGCCAAAGAACACTCTTGCGAGCTGGCTACTACTGGCGGGTCCCTGCAT 98
27 rGluValHISpHeArTyPPhenTyAsnArgThrSerLyAsrGysGluT 44
277      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
99 GGGAAATGACGACGAGGTATTCTATATATCGTACATCCATGGCCTGTGAGA 148
44 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLyLeuLy 60
149 CTTTCCAGTACGGCGCTGCGATGGGCAACGGTAACAACCTTCGTACAGAA 198
61 Ileglu.....ArgGluVal..... 65
199 AAGGAGTGTCTGCAGACCTCCGAACTGTGGCGCGCTGAAACTGTTCGA 248
56 AlaCysValAlaLy 70
249 AAGTGTGTTAGCAAAA 264

seq_name: /cgnl_7/prodata/1/ina/5A_COMB.seq:US-08-321-658B-17

seq_documentation_block:
; Sequence 17, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITIO
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5576294o No. 5576294disk of No. 5576294t
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693,210-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
;

```

```

; ORGANISM: synthetic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
; US-08-321-658B-17

alignment_scores:
  Quality: 127.00      Length: 63
  Ratio: 3.175        Gaps: 0
  Percent Similarity: 63.492    Percent Identity: 41.270

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-321-658B-17 ..
Align seg 1/1 to: US-08-321-658B-17 from: 1 to: 418

5  ILeAlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMe 21
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||
212 ATGCGTCGAGAGATTGGAGAGAGAAAGCCAGATTCTGCTTTTGGGAAGA 261
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||
21 tAsnPheGlySerCysTyrGluValHisPheAtgTyrPheTyrAsnArgT 38
   :: ||| ||| ::::::::::::::: ||| ||| ||| ||| ||| |||
262 AGATCTCGGAATATGTAAAGCTCGTATTATCAGGTATTTTATAACAATC 311
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||
38 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 54
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||
312 AGACAAACACAGTGTGAACGTTTCAAGTATGTTGATGTCAGGGCAATATG 361
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||
55 AsnAsnPheLysLeuLysIleGluArgGluValAlaCys 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 AACAAATTTTAAGACACTGGAAGAAATGCAAGAACATTTGT 400
   ::::::::::::::: ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgnl_7/ptodata/1/lna/5A_COMB.seq:US-08-321-658B-15

seq_documentation_block:
; Sequence 15, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBIT
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 5576294o No. 5576294disk of No. 5576294
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693.210-US

```

seq name: /csp1 7/prodata/1/ina/5A_COMB.seq:US-08-321-658B-15

FILE OF INVENTION: 18
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5576294o No. 5576294dsk of No. 5576294th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/321,658B
FILING DATE: 12-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 3693.210-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: gDNA
ORIGINAL SOURCE:

```

seq_documentation_block:
: Sequence 15: Application US/08321658B
:
: Patent No. 5576294
:
: GENERAL INFORMATION:
:
: APPLICANT: NO. 5576294ris, Panny
: APPLICANT: NO. 5576294ris, kjeld
: APPLICANT: Bjorn, Soren E
: APPLICANT: Petersen, Lars C
: APPLICANT: Olsen, Ole H
:
: INHIBITOR
:
: BROTASE
:
: TYPE
:
: KUNITZ
:
:

```

7
8
9
CORRESPONDENCE ADDRESS:
ADDRESS: NO. 55762940 NO. 5576294disk of No. 5576294th America, Inc.
empem. 405 Lexington Avenue, 64th Floor

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/321.658B
FILING DATE: 12-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 3693.210-US

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: synthetic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
US-08-321-658B-15

```

```

alignment_scores:
  Quality: 123.00      Length: 63
  Ratio: 3.075         Gaps: 0
  Percent Similarity: 63.492      Percent Identity: 39.683

```

```

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-321-658B-15 ..
Align seg 1/1 to: US-08-321-658B-15 from: 1 to: 418

5 IleAlaGluLysLeuCysGlyAspLeuLysAspProCysLysLeuAspMe 21
:::|||||::: ||| ||| |||:::
212 ATGGCTGAGAGATGGAGAGAGAGAGAGAGATTTCTGCTTTTGAAGA 261

21 tAsnPhcGlySerCysTyrGluValHisPheValPheSerGlyCysAsnGlyAsnLeu 38
::: ||| ||| :::::|||||:::
262 AGATCTGGAAATATGAAGCTCGTATTATCAGGTATTTTATAACAATC 311

38 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 54
:::|||||::: ||| :::::|||||:::
312 AGACAAACAGTGTGAAGCTTTCAAGTATGTTGGATCGAGGGCAATATG 361

55 AsnAsnPhcLysLeuLysLeuGluValHisPheValPheSerGlyCysAsnGlyAsnLeu 67
|||||::: ||| ::: |||
362 AACAAATTTTGAGACACTGGAAGAAATGCAAGAACATTTGT 400

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-147-710-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08147710
; Patent No. 5455338
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,710
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 93-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 979 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Placenta
; IMMEDIATE SOURCE:
; CLONE: J-2-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 39..746
US-08-147-710-1

```

```

alignment_scores:
  Quality: 121.00      Length: 70
  Ratio: 2.814         Gaps: 0
  Percent Similarity: 61.429      Percent Identity: 32.857

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-147-710-1 ..
Align seg 1/1 to: US-08-147-710-1 from: 1 to: 979

6 AlaGluLysLeuCysGlyAspLeuLysAspProCysLysLeuAspMetas 22
|||||::: |||::: ||| ||| |||:::
111 GCTCAGAGGCAACAGAGAAATAACGGGAGATGCTCTCTCGCCCTAGA 160

22 nPhcGlySerCysTyrGluValHisPheValPheSerGlyCysAsnGlyAsnLeu 39
|||||::: ||| :::::|||||:::
161 CTACGGACCTCGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 210

39 erLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsn 55
::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
211 CGCAGAGCTCGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260

56 AsnPhcLysLeuLysLeuGluValHisPheValPheSerGlyCysValAlaLysTyr 72
|||||::: |||::: |||::: |||::: |||::: |||::: |||:::
261 AATTCTACACCTGGGAGGCTTGGACGATGCTGTGGAGGATAGAAAA 310

72 SPOTPOArg 75
| |||:::
311 AGTTCCTCAA 320

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-458-090-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08458090
; Patent No. 5728674
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:

```

us-09-215-435-177_copy_25_99.rn1

Thu Nov 15 10:51:58 2001

```
; Sequence 1, Application US/08457887
; Patent No. 5914315
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kistler, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,887
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-14D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 979 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORGANISM: Homo sapiens
; TISSUE TYPE: Placenta
; IMMEDIATE SOURCE:
; CLONE: J-2-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 39..746
; US-08-458-090-1

alignment_scores:
  Quality: 121.00      Length: 70
  Ratio: 2.814        Gaps: 0
  Percent Similarity: 61.429      Percent Identity: 32.857

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-457-887-1 ..
Align seg 1/1 to: US-08-457-887-1 from: 1 to: 979

6 AlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAs 22
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
111 GCTCAGGAGCCCAACAGGAATAACGGGAGATCTCTCTCTGCCCTAGA 160

22 nPheGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrS 39
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 CTACGGACCTCGCGGCCCTACTTCTCGTTACTACTACGACAGGTACA 210

39 erLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsn 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 CGCAGAGCTGCCGCCAGTTCTGTACGGGGCTCGCAGGGCAACGCCAAC 260

56 AsnPhelLysLeuLysIleGluArgGluValAlaCysValAlaLysTyrLy 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 AATTCTACACCTGGGAGGCTTCGACGATGCTTGTGAGGATAGAAA 310

72 sProArg 75
|||||:
311 AGTTCCCAAA 320

seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-457-887-1
seq_documentation_block:
```



```

FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 77..235
;
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 236..409
;
US-08-321-658B-11

alignment_scores:
  Quality: 119.00      Length: 63
  Ratio: 3.132        Gaps: 0
  Percent Similarity: 60.317  Percent Identity: 39.683

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-321-658B-11 ..
Align seg 1/1 to: US-08-321-658B-13 from: 1 to: 418

5 IleAlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMe 21
:::|||||:::
212 ATGGCTGAGAGATTGGAGAGAGAAAGCCAGATTCTGCTTTTGAAGA 261
:::|||||:::
21 tAsnPhcGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgT 38
:::|||||:::
262 AGATCTGGAATATGTAAGCTCTATATATACAGTATTTTATAACAATC 311
:::|||||:::
38 hrSerLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeu 54
:::|||||:::
312 AGACAAAACAGTGAAGCTTTCAAGTATGTCGATGCTGGGCAATATG 361
:::|||||:::
55 AsnAsnPhcLysLeuLysIleGluArgGluValAlaCys 67
|||||:::
362 AACAATTTTGAGACACTCGAAGATGCAAGACATTTGT 400
|||||:::

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-321-658B-11

seq_documentation_block:
; Sequence 11, Application US/08321658B
; Patent No. 5576294
; GENERAL INFORMATION:
; APPLICANT: No. 5576294ris, Fanny
; APPLICANT: No. 5576294ris, Kjeld
; APPLICANT: Bjorn, Soren E
; APPLICANT: Petersen, Lars C
; APPLICANT: Olsen, Ole H
; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR
; TITLE OF INVENTION: VARIANT
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5576294o No. 5576294disk of No. 5576294th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.658B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3693.210-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: synthetic/human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..409

```


APPLICANT: Rasmussen, Jesper
 TITLE OF INVENTION: Anticoagulant Protein
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 53127360 No. 5312736dsk Of No. 5312736dsk of No. 5312736th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6200
 CITY: New York
 STATE: New York
 COUNTRY: United States of America

474 GTGGATCGCTGGCAATATGAACAATTTTGAGACACTCGAGGAATGCAAG 523

65 ValAlaCys 67

|||

524 AACATTTGT 532

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-446-646-8

seq_documentation_block:

; Sequence 8, Application US/08446646

; Patent No. 5726038

; GENERAL INFORMATION:

; APPLICANT: Christiansen, Lars

; APPLICANT: Petersen, Jens G.

; TITLE OF INVENTION: A DNA Construct Encoding the YAP3 Signal

; TITLE OF INVENTION: Peptide

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 57260380 No. 5726038disk of No. 5726038th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,646

; FILING DATE: 25-MAY-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 3987.204-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123

; TELEFAX: 212-878-9655

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 928 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cdna

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 8..919

; FEATURE:

; NAME/KEY: sig_peptide

; LOCATION: 8..91

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 92..919

US-08-446-646-8

alignment_scores:

Quality: 116.00

Ratio: 3.515

Percent Similarity: 62.264

Length: 53

Gaps: 0

Percent Identity: 43.396

alignment_block:

US-09-215-435-177_COPY_25_99 x US-08-446-646-8 ..

Align seg 1/1 to: US-08-446-646-8 from: 1 to: 928

15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

374 GATTTCTGCTTTTGGAAAGAGATCTCGAATATGTCGAGGTTATATTAC 423

31 eAryrTyPheTyAsnArgThrSerLysArgCysGluThrPheValPheS 48

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

424 CAGGTATTTTATACATCATCAGACAAACAGTGTGAAGGTTCAAGTATG 473

48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64

::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

474 GTGGATCGCTGGCAATATGAACAATTTTGAGACACTCGAGGAATGCAAG 523

65 ValAlaCys 67

|||

524 AACATTTGT 532

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-321-658B-7

seq_documentation_block:

; Sequence 7, Application US/08321658B

; Patent No. 5576294

; GENERAL INFORMATION:

; APPLICANT: No. 5576294ris, Fanny

; APPLICANT: No. 5576294ris, Kjeld

; APPLICANT: Bjorn, Soren E

; APPLICANT: Petersen, Lars C

; APPLICANT: Olsen, Ole H

; TITLE OF INVENTION: A HUMAN KUNITZ-TYPE PROTEASE INHIBITOR

; TITLE OF INVENTION: VARIANT

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 55762940 No. 5576294disk of No. 5576294th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/321,658B

; FILING DATE: 12-OCT-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lowney Dr., Karen A.

; REGISTRATION NUMBER: 31,274

; REFERENCE/DOCKET NUMBER: 3693.210-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123

; TELEFAX: 212-878-9655

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 945 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cdna

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 365..538

US-08-321-658B-7

alignment_scores:

Quality: 116.00

Ratio: 3.515

Percent Similarity: 62.264

Length: 53

Gaps: 0

Percent Identity: 43.396

Thu Nov 15 10:51:58 2001

```

alignment_block:
US-09-215-435-177_copy_25_99 x US-08-321-658B-7  ..
Align seg 1/1 to: US-08-321-658B-7 from: 1 to: 945

15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
371 GATTCTGCTTTTGGAGAGAGATCCTCGAATATGTCGAGGTTATATTAC 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 CAGGTATTTTATAACAATCAGACACAAACAGTGTGACGTTTCAAGTATG 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
471 GTGGATCCCTGGGCAATATGAACAATTTTGACACACTGGAGCAATGCAAG 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 ValAlaCys 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
521 ACAATTGT 529

```

OM of: US-09-215-435-177_COPY_25_99 to: Pending_Patents_NA_Main.* out_format : pfs
Date: Nov 15, 2001 3:55 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001.141105-2173/app_query.fasta_1.1519
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=nrpm
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435@cgnl_1.12482 -NCPU=6 -ICPU=3 -LONGLOG
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-177_COPY_25_99

Query length: 75

Database: Pending_Patents_NA_Main.*

Database sequences: 17159718

Database length: 173266264

Search time (sec): 11750.060000

```
score_list:          Strd Orig      zScore      EScore Len ! Documentation ...
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-5799 + 419.00 1013.36 4.2e-47 499
/cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-5799 + 419.00 1013.36 4.2e-47 49
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-76 + 419.00 1012.84 4.5e-47 526
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-303 + 409.00 988.06 1.1e-45 539
/cgnl_7/ptodata/1/pna/US6006_COMB.seq:US-60-069-957-136 + 409.00 988.06 1.1e-45 539
/cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-516-448-1671 + 368.00 891.74 2.5e-40 34
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-833-381-1671 + 368.00 891.74 2.5e-40 349
/cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-843-620-654 + 368.00 891.74 2.5e-40 349
/cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-188-162-1870 - 335.00 805.64 1.6e-35 58
/cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-148 + 307.00 744.54 4.0e-32 271
/cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-208-69337 + 251.00 606.72 1.9e-24 2
/cgnl_7/ptodata/1/pna/US091_COMB.seq:US-09-131-380-189 + 210.00 509.11 5.2e-19 208
/cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-499-23002 + 210.00 509.11 5.2e-19 2
/cgnl_7/ptodata/1/pna/US6005_COMB.seq:US-60-055-703-189 + 210.00 509.11 5.2e-19 208
/cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-480-902-6802 + 134.00 316.03 2.9e-08 404
/cgnl_7/ptodata/1/pna/US099_COMB.seq:US-09-960-352-6802 + 134.00 316.03 2.9e-08 404
/cgnl_7/ptodata/1/pna/US099_COMB.seq:US-09-480-902-13925 + 134.00 315.67 3.1e-08 41
/cgnl_7/ptodata/1/pna/US099_COMB.seq:US-09-960-352-13925 + 134.00 315.67 3.1e-08 41
/cgnl_7/ptodata/1/pna/US099_COMB.seq:PCT-US00-14462A-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US099_COMB.seq:PCT-US99-01313-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-435-810-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-373-658-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-373-658B-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-373-658C-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US0917_COMB.seq:US-60-171-503-23 + 133.00 267.51 1.5e-05 425
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-167-217-6889 + 131.00 307.64 8.6e-08 44
/cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-191-637-6788 + 131.00 305.87 1.1e-07 53
/cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-191-637-24855 + 131.00 291.05 7.2e-07 53
/cgnl_7/ptodata/1/pna/US6014_COMB.seq:US-60-142-896-264 - 131.00 290.52 7.7e-07 253
/cgnl_7/ptodata/1/pna/US6014_COMB.seq:US-60-191-637-6787 + 131.00 290.49 7.7e-07 25
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-167-217-6888 - 131.00 289.20 9.1e-07 28
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-167-217-6885 - 131.00 284.45 1.7e-06 46
/cgnl_7/ptodata/1/pna/US6015_COMB.seq:US-60-150-584-381 + 131.00 277.87 3.9e-06 908
/cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-528-237A-859 - 131.00 247.60 0.0002 19
/cgnl_7/ptodata/1/pna/US088_COMB.seq:US-08-849-406-3 + 130.00 309.66 6.6e-08 285
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-160-203-3184 + 129.50 302.23 1.7e-07 53
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-160-190-812 + 129.50 302.21 1.7e-07 53
/cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-160-203-2095 + 129.50 302.21 1.7e-07 53
/cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-192-739-1393 + 129.50 299.02 2.6e-07 73
/cgnl_7/ptodata/1/pna/US6014_COMB.seq:US-60-143-992-158 - 129.00 285.37 1.5e-06 259
```

/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-331-793-28 + 128.50 301.66 1.8e-07 44
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-235-076-30403 + 128.50 300.19 2.2e-07
/cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-289-768-37096 + 128.50 300.19 2.2e-07
/cgnl_7/ptodata/1/pna/US093_COMB.seq:US-09-332-782-30403 + 128.50 300.19 2.2e-07
/cgnl_7/ptodata/1/pna/US097B_COMB.seq:US-09-737-223-30403 + 128.50 300.19 2.2e-07

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-5799

seq_documentation_block:

Sequence 5799, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejanin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 5799
LENGTH: 499
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 22..318
NAME/KEY: sig_peptide
LOCATION: 22..93
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 4.59999990463257
OTHER INFORMATION: seq FFIFCSLNTLLG/GV
US-09-834-366-5799

alignment_scores:

Quality: 419.00 Length: 75
Ratio: 5.587 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-177_COPY_25_99 x US-09-834-366-5799 ..

Align seg 1/1 to: US-09-834-366-5799 from: 1 to: 499

```
1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProcy 17
|||||
94 GGTGTTAATAAATTCGCGAGAGATATGTGGAGACCTCAAGATCCCTG 143

17 sLysLeuAspMetAsnPhcGlySerCystyrGluValHisPheArgTyrP 34
|||||
144 CAATTTGGACATGAATTTTGGAGCTGCTATGAAGTWCATTTAGATATT 193

34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
|||||
194 TCTACAAACAGAACCTCCAAAAGATGTGAACATTTTCTCTCCGGCTGT 243

51 AsnGlyAsnLeuAsnAsnPhelysLeuLysIleGluArgGluValAlaCy 67
|||||
244 ATGGCAACTTAAACACTTCAAGCTTAAATAGAACGTGAAGTAGCCTG 293

67 sValAlaLysTyrLysProProArg 75
|||||
294 TGTGTGCAAAATAACAAACACCGAGG 318
```

seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-5799

seq_documentation_block:

Sequence 5799, Application US/60197873

Thu Nov 15 10:51:58 2001

```

; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Joberit, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 5799
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..318
; NAME/KEY: sig_peptide
; LOCATION: 22..93
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.59999990463257
; OTHER INFORMATION: seq FFIFCSLNTLLG/GV
; US-60-197-873-5799

alignment_scores:
  Quality: 419.00      Length: 75
  Ratio: 5.587        Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-215-435-177_COPY_25_99 x US-60-197-873-5799  ..
  Align seg 1/1 to: US-60-197-873-5799 from: 1 to: 499

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCy 17
|||||
94 GGTGTTAATAAAATTCGGAGAGATATGTGGAGACCTCAAGATCCCTG 143
|||||
17 slsLeuAspMetAsnPhcGlySerCysTyrGluValHisPheArgTyrP 34
|||||
144 CAAATGGACATGAATTTGGAGCTGCTATGAAGTTCACCTTTAGATATT 193
|||||
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
|||||
194 TCTACAACAGAACCTCCAAAGATGTGAAGCTTTTGTCTTCTCCGCTGT 243
|||||
51 AsnGlyAsnLeuAsnAsnPhcLysLeuLysIleGluArgGluValAlaCy 67
|||||
244 AATGGCAACCTTAACAACCTCAAGCTTAAATAGAACGTGAAGTAGCCTG 293
|||||
67 sValAlaLysTyrLysProProArg 75
|||||
294 TGTGCAAAATACAAACACCGAGG 318

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-76

seq_documentation_block:
; Sequence 76, Application US/09215435A
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bougueleret Lydie
; TITLE OF INVENTION: Extended cDNAs
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-2-9

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert Aymeric,
; APPLICANT: Bougueleret Lydie
; TITLE OF INVENTION: Extended cDNAs
; FILE REFERENCE: GENSET.019A
; CURRENT APPLICATION NUMBER: US/09/215,435A
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069,957
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/074,121

```

EARLIER FILING DATE: 1998-2-9
EARLIER APPLICATION NUMBER: 60/081,563
EARLIER FILING DATE: 1998-4-13
EARLIER APPLICATION NUMBER: 60/096,116
EARLIER FILING DATE: 1998-8-10
NUMBER OF SEQ ID NOS: 519
SOFTWARE: Patent.pm
SEQ ID NO 303
LENGTH: 539
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 32..328
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 32..103
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 4.59999990463257
OTHER INFORMATION: seq FFIFCSLNTLLG/GV
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 508..513
FEATURE:
NAME/KEY: polyA_site
LOCATION: 528..539
US-09-215-435-303

alignment_scores:
Quality: 409.00 Length: 75
Ratio: 5.453 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.333
alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-215-435-303 ..
Align seg 1/1 to: US-09-215-435-303 from: 1 to: 539
1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
104 GGTGTTAAATAAAATTCGGAGAGATATGTTGGAGACCTCAAAGATCCCTG 153
17 sLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrP 34
154 CAAATTGGACATGAATTTTGGAGCTGCTATGAAGTTCACTTAGATATT 203
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
204 TCTACAACAGAACCTCCAAAGATGTGAAACTTTTGTCTCTCCAGCTGT 253
51 AsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCy 67
254 AATGGCAACCTTAACAACCTCAAGCTTAAATAGAACGTGAAGTAKCCTG 303
67 sValAlaLysTyrLysProProArg 75
304 TGTGCAAAATACAAACACCGAGG 328.

seq_name: /cgnl_7/ptdata/1/pna/us6006_COMB.seq:US-60-069-957-136
seq_documentation_block:
Sequence 136, Application US/60069957
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste Dumas Milne
APPLICANT: Duclert, Aymeric
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 381
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California

COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/069,957
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET.019PR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Testis
FEATURE:
NAME/KEY: CDS
LOCATION: 32..328
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 32..103
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.6
OTHER INFORMATION: seq FFIFCSLNTLLG/GV
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 508..513
FEATURE:
NAME/KEY: poly_a
LOCATION: 528..539
IDENTIFICATION METHOD: blastn
US-60-069-957-136

alignment_scores:
Quality: 409.00 Length: 75
Ratio: 5.453 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.333
alignment_block:
US-09-215-435-177_COPY_25_99 x US-60-069-957-136 ..
Align seg 1/1 to: US-60-069-957-136 from: 1 to: 539
1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCys 17
104 GGTGTTAAATAAAATTCGGAGAGATATGTTGGAGACCTCAAAGATCCCTG 153
17 sLysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrP 34
154 CAAATTGGACATGAATTTTGGAAAGCTGCTATGAAGTTCACTTAGATATT 203
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
204 TCTACAACAGAACCTCCAAAGATGTGAAACTTTTGTCTCTCCAGCTGT 253
51 AsnGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCy 67
254 AATGGCAACCTTAACAACCTCAAGCTTAAATAGAACGTGAAGTAKCCTG 303

Thu Nov 15 10:51:58 2001

us-09-215-435-177_copy_25_99.inpnm

```
67 sValAlaLysTyrIysProProArg 75
|||||
304 TCTTCAAAATCAAAACCCAGG 328

seq_name: /cgnl_7/ptodata/1/pna/US095A_COMB.seq:US-09-516-448-1671

seq_documentation_block:
; Sequence 1671, Application US/09516448
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/516,448
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(349)
; OTHER INFORMATION: n = A,T,C or G
US-09-516-448-1671

alignment_scores:
Quality: 368.00 Length: 68
Ratio: 5.662 Gaps: 0
Percent Similarity: 95.588

- alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-833-381-1671 ..
Align seg 1/1 to: US-09-833-381-1671 from: 1 to: 349
8 LysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheG1 24
|||||
23 AAGATATGTGGAGACCTCAAGATCCCTGCANATTGGACATGAATTTGG 72
|||||
24 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 41
|||||
73 AAGCTGCTATGAAGTTTCACTTTAGATATTTTCTACAAACAGAACCTCCANAA 122
|||||
41 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 57
|||||
123 GATGTGAAACTTTTGNCTTCTCCGGCTGTAATGGCAACCTTAACAACCTTC 172
|||||
58 LysLeuLysIleGluArgGluValAlaCysValAlaLysTyrIysProPr 74
|||||
173 AAGCTTAAATANAACGTGAAGTAGCCTGTGTTGCAAAATACAAACACC 222
|||||
74 oArg 75
|||||
223 GAGG 226

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-843-620-654

seq_documentation_block:
; Sequence 654, Application US/09843620
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2098-001
; CURRENT APPLICATION NUMBER: US/09/843,620
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/200,099
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 1135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 654
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(349)
; OTHER INFORMATION: n = A,T,C or G
US-09-843-620-654

alignment_scores:
Quality: 368.00 Length: 68
Ratio: 5.662 Gaps: 0
Percent Similarity: 95.588

- alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-843-620-654 ..
Align seg 1/1 to: US-09-516-448-1671 from: 1 to: 349
8 LysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheG1 24
|||||
23 AAGATATGTGGAGACCTCAAGATCCCTGCANATTGGACATGAATTTGG 72
|||||
24 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 41
|||||
73 AAGCTGCTATGAAGTTTCACTTTAGATATTTTCTACAAACAGAACCTCCANAA 122
|||||
41 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 57
|||||
123 GATGTGAAACTTTTGNCTTCTCCGGCTGTAATGGCAACCTTAACAACCTTC 172
|||||
58 LysLeuLysIleGluArgGluValAlaCysValAlaLysTyrIysProPr 74
|||||
173 AAGCTTAAATANAACGTGAAGTAGCCTGTGTTGCAAAATACAAACACC 222
|||||
74 oArg 75
|||||
223 GAGG 226

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-833-381-1671

seq_documentation_block:
; Sequence 1671, Application US/09833381
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 349
```


Align seg 1/1 to: US-09-843-620-654 from: 1 to: 349

```
8 LysileCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheG1 24
|||||
23 AGATATGTTGGAGACCTCAAGATCCTCGAATTCGACATGAATTTGG 72
|||||
24 ySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysA 41
|||||
73 AAGCTGCTATGAAGTTCACCTTTAGATATTTCTACACAGAACCTCCANAA 122
|||||
41 rgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhe 57
|||||
123 GATGTGAAACTTTTGNCTTCTCCGGCTGTAATGSCAACCTTAACAACCTTC 172
|||||
58 LysLeuLysTleGluArgGluValAlaCysValAlaLysTyrLysProPr 74
|||||
173 AAGCTTAAATAANACGTGAAGTAGCCTGTGTGCAAAATACAAACACC 222
|||||
74 oArg 75
|||||
223 GAGG 226
```

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-188-162-1870

```
seq_documentation_block:
; Sequence 1870, Application US/60188162
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: C1000518
; CURRENT APPLICATION NUMBER: US/60/188,162
; CURRENT FILING DATE: 2000-05-09
; NUMBER OF SEQ ID NOS: 5094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1870
; LENGTH: 585
; TYPE: DNA
; ORGANISM: HUMAN
US-60-188-162-1870
```

alignment_scores:
Quality: 335.00 Length: 60
Ratio: 5.583 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.333

alignment_block:
US-09-215-435-177_COPY_25_99 x US-60-188-162-1870/rev ..

Align seg 1/1 to reverse of: US-60-188-162-1870 from: 1 to: 585

```
15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValHisPh 31
|||||
256 GATCCCTGCAAAATTGACATGAATTTTGGAAAGCTGCTATGAAGTTCACIT 207
|||||
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
|||||
206 TAGATATTTCTACACAGAACCTCCAAAGAGATGTGAAACTTTTGTCTTCT 157
|||||
48 erCysGlyAsnGlyAsnLeuAsnAsnPheLysLeuLysTleGluArgGlu 64
|||||
156 CCAGCTGTATGCAACCTTAACAACCTTAAGCTTAAATAAGACGTGAA 107
|||||
65 ValAlaCysValAlaLysTyrLysProPro 74
|||||
106 GTAGCTGTGTTGCCAAATACAAACACCG 77
|||||
```

seq_name: /cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-148

seq_documentation_block:

```
; Sequence 148, Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; TITLE OF INVENTION: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,279
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 32..103
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.6
; OTHER INFORMATION: seq FFIFCSLNTLLLG/GV
US-08-905-279-148
```

alignment_scores:
Quality: 307.00 Length: 55
Ratio: 5.582 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.182

alignment_block:

US-09-215-435-177_COPY_25_99 x US-08-905-279-148 ..

Align seg 1/1 to: US-08-905-279-148 from: 1 to: 271

```
1 GlyValAsnLysTleAlaGluLysIleCysGlyAspLeuLysAspProcy 17
|||||
104 GGTGTTAATAAAATTCGGAGAAAGATATGTGGAGACCTCAAGATCCCTG 153
|||||
17 LysLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrP 34
|||||
154 CAAATTTGGACATGAATTTTGGAAAGCTGCTATGAAGTTCACCTTTAGATATT 203
|||||
34 heTyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCys 50
|||||
204 TCTACACAGAACCTCCAAAGAGATGTGAAACTTTTGTCTTCTCAGCTGT 253
|||||
51 AsnGlyAsnLeuAsn 55
|||||
254 AATGGCAACCTTAAC 268
```

```

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT FEMUR
; FILE REFERENCE: PZ-0013 US
; CURRENT APPLICATION NUMBER: US/09/131,380
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,705
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/074,278
; EARLIER FILING DATE: 1998-02-10
; NUMBER OF SEQ ID NOS: 2199
; SOFTWARE: PERL Program
; SEQ ID NO 189
; LENGTH: 208
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 109, 117, 156, 167
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE: -
; OTHER INFORMATION: 700292496H1
; US-09-131-380-189

```

```

alignment_scores:
  Quality: 210.00      Length: 67
  Ratio: 4.038         Gaps: 1
  Percent Similarity: 77.612      Percent Identity: 55.224

```

```

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-131-380-189 ..
Align seg 1/1 to: US-09-131-380-189 from: 1 to: 208

```

```

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 GTGGAGAGGTTGCTGCAAGGCTCTGCAAGAAATTCACAGATCCTTGTC 51
18 sLeuAspMetAsnPhGlySerCysTyrGluValHisPheArgTyrPheT 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
52 TTTGGAATGGATCCTGGCAGCTGCTATGAAAGCCATTACAGATTTTCT 101
35 yrAsnArgThrSer.LysArgCysGluThrPheValPheSerGlyCysAs 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
102 ACAACCAACCTCCACCAGAGTGTACAGATTTTCGTGTTCAAGTTGCCAA 151
51 nGlyAsnLeuAsnPhGlySerCysTyrGluValHisPheArgTyrPheT 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 TGTNACCGTAAATTAATTCATGCTTAAATAGACTCTGACGTAATGTGT 200

```

seq_name: /cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-499-23002

```

seq_documentation_block:
; Sequence 23002, Application US/09540499
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MUSCULOSKELETAL SYSTEM TISSUE
; CURRENT APPLICATION NUMBER: US/09/540,499
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 29520
; SOFTWARE: PERL Program
; SEQ ID NO 23002
; LENGTH: 208
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature

```

seq_name: /cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-208-69337

```

seq_documentation_block:
; Sequence 69337, Application US/09540208
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 69337
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: rat00103589
; NAME/KEY: unsure
; LOCATION: 22, 255
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-208-69337

```

```

alignment_scores:
  Quality: 251.00      Length: 71
  Ratio: 4.328         Gaps: 0
  Percent Similarity: 81.690      Percent Identity: 59.155

```

```

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-540-208-69337 ..
Align seg 1/1 to: US-09-540-208-69337 from: 1 to: 282

```

```

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 GTGGAGAGGTTGCTGCAAGGCTCTGCAAGAAATTCACAGATCCTTGTC 50
18 sLeuAspMetAsnPhGlySerCysTyrGluValHisPheArgTyrPheT 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 TTTGGAATGGATCCTGGCAGCTGCTATGAAAGCCATTACAGATTTTCT 100
35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 ACAACCAACCTCCAAAAGTGTACAGATTTTCGTGTTCAAGTTGCCAAT 150
52 GlyAsnLeuAsnPhGlySerCysTyrGluValHisPheArgTyrPheT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 GGTAACCGTAAATTAATTCATGCTTAAATAGACTCTGACGTAATGTGTA 200
68 lAlaLysTyrLys 72
201 TGAAGAATACAAG 213

```

seq_name: /cgnl_7/ptodata/1/pna/US091_COMB.seq:US-09-131-380-189

```

seq_documentation_block:
; Sequence 189, Application US/09131380
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Garrow, Bonnie L.
; APPLICANT: Klemm, Jull D.

```

239 GATTCTGCCAGCTGGACTACTCACAAGGGCCTTGCCTGGGGCTGTTCAA 288

Thu Nov 15 10:51:58 2001

```
31 eArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValPheS 48
   |||||
289 GAGGTATTCTATACGGGTACATCCATGGCCTGTGAGACCTTCTTATATG 338
   |||||
48 erGlyCysAsnGlyAsnLeuAsnAsnPhe 57
   ::|||
339 GCGGCTGCATGGGGAATGCCAACAACTTC 367
```


us-09-215-435-177_copy_25_99.rnpn

Thu Nov 15 10:51:59 2001

```

11 GlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCystY 27
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 GCGCCCAAGCCTGACTTCTGCTTCCTCGAGGAGGATCCGGGATTGCGG 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 rGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysGluT 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 CGGTTATATTACGGCTTATTCTATATAAACCAGACTAAGCAATGTGAGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLys 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 GGTTCAGTATGTTGCTGCTAGTAAATATGACAACTTCGAGACTCTA 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 IleGluArgGluValAlaLys 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 GAAGAGTGTGAAGAACATATGT 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-84

seq_documentation_block:
; Sequence 84, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
;
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
;
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
;
; US-08-849-406A-84

```

```

alignment_scores:
  Quality: 105.00      Length: 57
  Ratio: 3.387        Gaps: 0
  Percent Similarity: 54.386      Percent Identity: 36.842

alignment_block:
US-09-215-435-177_copy_25_99 x US-08-849-406A-84

```

```

44 hrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPhelLysLeuLys 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 CTTTCAGTACGGCGCTGCATGGCAACGGTAACAACCTTCGTACACAGAA 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 IleGlu.....ArgGluVal..... 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 AAGGAGTGTCTGCAGACCTGCCGAACCTGTGGCGCGCGCTGAACACTGTTGA 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 AlaCysValAlaLys 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 AAGTTGTTTAGCAAA 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-78

seq_documentation_block:
; Sequence 78, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
;
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
;
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 78:
;
; US-08-849-406A-78

```

```

alignment_scores:
  Quality: 117.00      Length: 57
  Ratio: 3.441        Gaps: 0
  Percent Similarity: 59.649      Percent Identity: 42.105

alignment_block:
US-09-215-435-177_copy_25_99 x US-08-849-406A-78
Align seg 1/1 to: US-08-849-406A-78 from: 1 to: 249

```

[illegible]

```

seq_name: /cgml_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-11942

seq_documentation_block:
; Sequence 11942, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FROM CDNA LIBRARIES OF FETAL LIVER-SPLEEN AND INFANT BRAIN
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11942
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(538)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-11942

alignment_scores:
Quality: 99.50 Length: 59
Ratio: 3.109 Gaps: 2
Percent Similarity: 54.237 Percent Identity: 38.983

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-922-340-11942 ..
Align seg 1/1 to: US-09-922-340-11942 from: 1 to: 538

10 CysGlyAspLeuLysAspProCysLysLeuAspMetAsnPheGlySerCy 26
||| ||||| |||
331 TGGGTGGACCTGCCAGACACA.....GGACTCTG 359

26 sTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLysArgCysG 43
| ||| ||||| ||||| ||||| |||||
360 CAAGGAGAGCATCCGCGCTGGTACTACAAACCCCTTCAGTGAACACTGCG 409

43 luThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeu 59
|||||:|||||:|||||:|||||:|||||:
410 CCCGCTTTACCTATGCTGGTGGTGGTATGCGAACAGAACAACTTT..... 453

60 LysIleGluArgGluValAlaCysVal 68

```

Thu Nov 15 10:51:59 2001

us-09-215-435-177_copy_25_99.rnpn

188 TTAAA.....TCGCCGAAGATTGCATCGCTACCTCGGTGGCGCC 228
73 o 73
229 G 229

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-1

seq_documentation_block:
; Sequence 1, Application US/08849406A

; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.

; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ

; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.

; STREET: 624 Ninth Street, N.W. Suite 300

; CITY: Washington

; STATE: District of Columbia

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/849,406A

; FILING DATE: 21-Jul-1999

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/16349

; FILING DATE: 15-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Cooper, Iver P.

; REGISTRATION NUMBER: 28,005

; REFERENCE/DOCKET NUMBER: LEY-1A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-628-5197

; TELEFAX: 202-737-3528

; INFORMATION FOR SEQ ID NO: 1:

; LENGTH: 276 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-849-406A-1

alignment_scores:
Quality: 98.50 Length: 51
Ratio: 3.078 Gaps: 2
Percent Similarity: 62.745 Percent Identity: 43.137

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-849-406A-1 ..

Align seg 1/1 to: US-08-849-406A-1 from: 1 to: 276

24 GlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerLy 40
||| |||
88 GGGCCCTGCAAGCGCGCATCATCGCTATTTCTACATGTAAGCAGG 137

40 sArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsp 57
||| |||
138 CCTGTCCAGACCTTTGTATACGGTGGTTCGCGTGTAAAGCGTAACAACT 187

57 helYLeuYslyIleGluArgGlu.ValAlaCysValAlaLysTyrLysPr 73
||| |||
||| |||
||| |||

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-149

seq_documentation_block:
; Sequence 149, Application US/09971392

; GENERAL INFORMATION:

; APPLICANT: Peterson, David P.

; APPLICANT: Pearson, Cecilia I.

; APPLICANT: Cocks, Benjamin G.

; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION

; FILE REFERENCE: PA-0029 US

; CURRENT APPLICATION NUMBER: US/09/971,392

; CURRENT FILING DATE: 2001-10-03

; PRIOR APPLICATION NUMBER: 60/237,652

; PRIOR FILING DATE: 2000-10-03

; NUMBER OF SEQ ID NOS: 260

; SOFTWARE: PERL Program

; SEQ ID NO 149

; LENGTH: 6777

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Template ID: 235164.25

US-09-971-392-149

alignment_scores:
Quality: 92.50 Length: 68
Ratio: 2.312 Gaps: 2
Percent Similarity: 58.824 Percent Identity: 26.471

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-971-392-149 ..

Align seg 1/1 to: US-09-971-392-149 from: 1 to: 6777

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
|||
982 GTGGAAGAGGTGGTTCGAGAGGTGCTCTGCAACAGCCGAGAGC 1026

18 sleuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 35
||| |||
1027GGCGCGTCCGAGCAATGATCTCCGCGTGGTACT 1050

35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
::: |||
1061 TTGATGTGACTGAAGGAAGTGTGCCCATCTTTTACGGCGGATGTGGC 1110

52 GlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCysVa 68
||| |||
1111 GGCAACCGGAACAACCTTGACACAGAAGAGTACTGTCAT 1148

68 lAla 69

::: |||

1149 GGCC 1152

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-9347

seq_documentation_block:
; Sequence 9347, Application US/09922340

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained

; FROM INVENTION: From CDNA Libraries Of Fetal Liver-Spleen and Infant Brain

; FILE REFERENCE: 20411-726CON2

; CURRENT APPLICATION NUMBER: US/09/922,340

; CURRENT FILING DATE: 2001-08-03

REFLECTION NUMBER: 101/0000/10040

```

; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-08-849-406A-73

```

```

alignment_scores:
  Quality: 78.00      Length: 61
  Ratio: 2.294       Gaps: 0
  Percent Similarity: 55.738      Percent Identity: 27.869

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-849-406A-73 ..
Align seg 1/1 to: US-08-849-406A-73 from: 1 to: 444

7 GlutylSerCysGlyAspLeuAspProCysLeuAspMetAsnPh 23
|||||
241 GAAGTGTCTTCCCTGGACACAGAGAGCGCTGTAACCTGCCAATCCTCAG 290
|||||
23 eGlySerCysGlyGluValHisPheArgTyrPheTyrAsnArgThrSerL 40
|||||
291 AGGTCATGATGCTTCTTCCCAAGATGGCTTCGAGCGTGTAAAG 340
|||||
40 ysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsn 56
|||||
341 GTAAGTGGCTGTTGTTCCCATACGGTGGTGTCAAGGTAACGGTAACAA 390
|||||
57 PheLysLeuLysIleGluArgGluValAlaCys 67
|||||
391 TTCTACTGTAGAGAGAGTGTAGAGACTACTGT 423

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-11256

```

seq_documentation_block:
; Sequence 11256, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11256
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-340-11256

```

```

alignment_scores:
  Quality: 76.00      Length: 34
  Ratio: 3.455       Gaps: 1
  Percent Similarity: 64.706      Percent Identity: 44.118

```

```

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-922-340-11256 ..
Align seg 1/1 to: US-09-922-340-11256 from: 1 to: 401

35 TyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 51
|||||
2 TACAACCCCTTCAGTGAACACTGCGCCCGCTTACCTATGGTGGTTGTTA 51
|||||
51 nGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 68
|||||
52 TGGCAACAACAACAACACTT.....GAGGAAGAGACGACGAGTGC 89
|||||
68 al 68
::
90 TC 91

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-8534

```

seq_documentation_block:
; Sequence 8534, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8534
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(418)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-8534

```

```

alignment_scores:
  Quality: 75.00      Length: 33
  Ratio: 3.409       Gaps: 0
  Percent Similarity: 66.667      Percent Identity: 42.424

```

```

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-922-340-8534 ..
Align seg 1/1 to: US-09-922-340-8534 from: 1 to: 418

35 TyrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAs 51
|||||
2 TACAACCCCTTCAGTGAACACTGCGCCCGCTTACCTATGGTGGTTGTTA 51
|||||
51 nGlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 67
|||||
52 TGGCAACAACAACAACACTTGGAGGAAAGCAGCAGTGCCTCGAGTCTGT 100
|||||

```

seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-71

```

seq_documentation_block:
; Sequence 71, Application US/08849406A
; APPLICANT: LEV, Arthur C.
; LADNER, Robert C.
; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William

```

; KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
;
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
;
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8584 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 71:
US-08-849-406A-71

alignment_scores:
Quality: 74.00 Length: 61
Ratio: 2.242 Gaps: 0
Percent Similarity: 54.098 Percent Identity: 27.869

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-849-406A-71 ..
Align seg 1/1 to: US-08-849-406A-71 from: 1 to: 8584
7 GluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPh 23
1185 GAAGGTGTTCTTCCTGGACAGAGAGCTGTTGTAACCTGCCAATCGTCAG 1234
23 eGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerL 40
1235 AGTCCATGCAATGCTTCTTCCCAAGATGGGCTTTCGACGCTGTTAAGG 1284
40 ysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsn 56
1285 GTAAGTGCCTGTTGTTCCCATGAGGTGTTGTCAGGTAACGTAACAAG 1334
57 PheLysLeuLysIleGluArgGluValAlaCys 67
1335 TTCTACTCTGAGGAGGAGGTGTCAGAGTACTGT 1367
seq_name: /cgnl_7/ptodata/1/pna/US08_NEW_COMB.seq:US-08-849-406A-75
seq_documentation_block:
; Sequence 75, Application US/08849406A
; APPLICANT: LEY, Arthur C.
; LADNER, Robert C.

; GUTERMAN, Sonia K.
; ROBERTS, Bruce L.
; MARKLAND, William
; KENT, Rachel B.
;
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
;
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20001
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,406A
; FILING DATE: 21-Jul-1999
; CLASSIFICATION: 530
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16349
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
;
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-08-849-406A-75

alignment_scores:
Quality: 74.00 Length: 61
Ratio: 2.242 Gaps: 0
Percent Similarity: 54.098 Percent Identity: 27.869

alignment_block:
US-09-215-435-177_COPY_25_99 x US-08-849-406A-75 ..
Align seg 1/1 to: US-08-849-406A-75 from: 1 to: 8590
7 GluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAsnPh 23
1185 GAAGGTGTTCTTCCTGGACAGAGAGCTGTTGTAACCTGCCAATCGTCAG 1234
23 eGlySerCysTyrGluValHisPheArgTyrPheTyrAsnArgThrSerL 40
1235 AGTCCATGCAATGCTTCTTCCCAAGATGGGCTTTCGACGCTGTTAAGG 1284
40 ysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsnAsn 56
1285 GTAAGTGCCTGTTGTTCCCATGAGGTGTTGTCAGGTAACGTAACAAG 1334
57 PheLysLeuLysIleGluArgGluValAlaCys 67
1335 TTCTACTCTGAGGAGGAGGTGTCAGAGTACTGT 1367
seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-830-446-15
seq_documentation_block:

Thu Nov 15 10:51:59 2001

```
; Sequence 15, Application US/09830446
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.; OOMEN, Raymond P; and WANG, Joe
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there
; FILE REFERENCE: 77813-51
; CURRENT APPLICATION NUMBER: US/09/830,446
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/106034
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106039
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/10604
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106044
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 60/106072
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106073
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106074
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106087
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/106587
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/106588
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107089
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107034
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: US 60/107035
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2194)
; US-09-830-446-15

alignment_scores:
      Quality: 68.00      Length: 58
      Ratio: 2.061      Gaps: 2
Percent Similarity: 56.897 Percent Identity: 32.759

alignment_block:
US-09-215-435-177_COPY_25_99 x US-09-830-446-15 ..
Align seg 1/1 to: US-09-830-446-15 from: 1 to: 2300

23 PheGly.....SerCysTyrGluValHisPheArgTyrPheTyr... 35
|||||.....:|||||:.....:|||||:.....:
1522 TTGGAGGCGCTATTAAAGTGCCCTTAGATCAAGATCGAGATCTGTCATTG 1571
|||||.....AsnArgThrSerLysArgCysGluThrPheValPheS 48
|||||:|||||:.....:|||||:.....:
1572 TCTTGAATCCTAGAAATCATGGCTGAGGATACGCAACAAGCTTCGTTTTTT 1621
|||||.....:|||||:.....:|||||:.....:
48 erGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGlu 64
:|||||:|||||:.....:|||||:.....:
1622 GTAGGCAACCGGTCCTTACCAACTATCAAAATCTATATCCAGAAGAAC 1671
|||||.....:|||||:.....:
65 ValAlaCysValAlaLysTyrLys 72
||| |||||||:.....:
1672 AGGAACGTGAACCAAAATATCGA 1695
```


further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GACAGAGAGAAGATCAAGACGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GACAGAGATCTCGATCTGAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

```

FEATURES
source
Location/Qualifiers
1..611
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1912158"
/db_xref="MGD:MGI:1925489"
/clone="9230105115"
/sex="male"
/tissue_type="epididymis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

```

```
CDS
/dev_stages=adult
24.503
/note="putative"
/codon_start=1
/protein_id="BAB32069.1"
/protein_acc="12950869"
```

```

/db_xref="GI:12686069
/translation="MPTKLGFLGLSLCSLSPVLGVERLANVLCCKDYNDPCLL
VEPGSCYEVHFRFFYNQAKQCIQIFLTGCGNGLNNEFKLIDCDVTHETYNKPLP

```

polyA_signal

```
polyA_site      611
                /note="putative"
                152 145 140
```

ORIGIN	1/4 B	152 C	143 S	140 C
BASE COUNT				

alignment_scores:	
Quality:	269.00
Length:	74

```

Quantity: 4.270      Gaps: 0
Ratio: 4.270
Percent Similarity: 85.135
Percent Identity: 60.811

```

```
alignment_block:
US-09-215-435-177_COPY_25_99 x AK020315 ..
```

Align seg 1/1 to: AK020315 from: 1 to: 611

1 GlyValAsnLysIleAlaGluLysIleCysGlyAspLeuLysaspProC 17
||||| :||| ||| :|||
2 CAGATGGTGCCTCAACCGCTACTATTCTGTCAAAGACTACAACGATCCCTG 145

96 GGTTGATTAGAGGGCTACCGATCAACTTCCGCCTCCTCATTCCT
17 slysLeuaspmetasnPhcylSerCystStyrGlualHispheArgTyrP 34

146 TCCTGTTGGATGGAGCCTGGCAGCTGCTATGAAGTACATTTAGATTTT 195

34 heTyRAsnArgThrSerLysArgCysIuThrPheValPheSerGlyCys 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 mTmRCACACAAAGCGCCAAACAGTGTCAGATTTCTGTCACTGGCTGC 245

196 TCACACCAATGCTCCAGTGGTGTCTC
51 AsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArgGluValAlaCy 67

246 AATGGCAACCTTAATAACTTCAAGCTGAAAATAGACTGTGATGTAAC TTG 295

67 svaAlaLysTyrLysProPro 74
- | ||||| ||
206 TCACCAAAACATACAAAAACCT 317

[illegible]

```

51  AsnGlyAsnLeuAsnAsnPhenylsLeuLysIleGluuArgGluuValAlaCy 67
|||||
282  AATGGCAACCTTAAACAACCTTCAAGCTTAAATAGAACCTGAAGTAGCCTG 233
|||||

67  sValAlaLysTyrLysProProArg 75
|||||

232  TGTGTGCANAATACAAACACCGAGG 208
|||||

seq_name: gb_htc:AK020315 .
seq_documentation_block:
ACUUS AK020315 611 bp mRNA HTC 08-FEB-2001
DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
library, clone:9230105I15, full insert sequence.

```

ACCESSION	AK020315
VERSION	AK020315.1
ENTRY	GI:12860868
FEATURES	CAP TRAPPER
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
SOURCE	clone:lib-Riken full-length enriched mouse cDNA library
	clone:523010515.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE	1 (sites)
AUTHORS	Carninci, P. and Hayashizaki, Y.
	Gene cloning

TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE
Normalization and subtraction of cap trapped cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE	20499374	3 (sites)	Shibata K.	Ttoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
REFERENCE				
ABSTRACT				
ANNOTATION				
REFERENCES				
REMARKS				

ADIRONS

Shibata, K., Ito, M.,
Konno, H., Akiyama, J.,
Kikuchi, N., Ishii, Y.,
Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Nakamura, S., Hazama, M., Nishine, T.,
Matsumoto, H.
Sakaguchi, S., Ikegami, T.

Harada, A., Yamamoto, K., Matsumoto, K., Ogasawara, K., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Iizawa, Y., and Uvachizaki, Y.

TITLE	Author(s)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,I. and Nakashizuka,T.

JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913

4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and
FANTOM Consortium.

TITLE	FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE CDNA COLLECTION
NATURE 409, 685-690 (2001)	
REFERENCE	5 (bases 1 to 611)

AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Aral, A.,
Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
Hiramoto, K., Hiraoaka, T., Hori, F., Iwata, N., Katayama, M.,
Kawaguchi, M., Matsui, M., Nakamura, M., Ohtsuka, T., Shimizu, Y.,
Suzuki, Y., Taniguchi, M., Teramitsu, R., Ueda, Y., Watanabe, Y.

Hanayama, I., Hara, K., Maeda, Y., Matsuyama, T.,
Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
Nakagaki, Y., Nishimura, Y., Okazaki, Y., Shimizu, T.,
Shinozaki, T., Sugita, M., Takahashi, Y., Tanaka, A.,
Takahashi, K., Takemoto, T., Terada, T., Uchida, T.,
Uchiyama, Y., Yamashita, Y., Yanai, H., Yoshida, K.,
Yoshida, M., Yoshida, T., Yoshida, Y., Yokoyama, C.,
Yokoyama, T., Yokozaki, T., Yonemitsu, S.,
Yoshioka, S., Yoshioka, T., Yoshizawa, T.,
Zhang, L.

Miyazaki, A., Nishii, K., Nomura, K., Numazaki, Y., Ogasawara, T., Oka, C., Sano, H., Sakai, K., Saito, R., Saito, H., Shibaata, Y., Shinagawa, A., Shiraki, T., Sasaki, D., Shibaata, K., Shibaata, Y., Shibaata, Y., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagawa, A., Takanashi, F., Tanaka, T., Tejlina, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

[illegible]

Physical and Chemical Research (RIKEN), Exploratory Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 227-0216, Japan

Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

```

seq_name: gb_gss32:A2763929
seq_documentation_block:
LOCUS A2763929 570 bp DNA GSS 16-FEB-2001
DEFINITION IM0559C11R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0559C11 R, DNA sequence.
ACCESSION A2763929
VERSION A2763929.1 GI:12878349
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 570)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0559 row: C column: 11
Seq primer: CACACAGGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 570.
FEATURES
source
location/Qualifiers
1..570
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0559C11"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/lab_host="Male"
/sex="Male"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 179 a 123 c 126 g 142 t
ORIGIN

alignment_scores:
Quality: 235.00 Length: 60
Ratio: 4.519 Gaps: 0
Percent Similarity: 86.667 Percent Identity: 65.000

alignment_block:
US-09-215-435-177_COPY_25_99 x A2763929/rev ..
Align seg 1/1 to reverse of: A2763929 from: 1 to: 570

15 AspProCysLysLeuAspMetAsnPheGlySerCysTyrgluValHisph 31
||||| ||||||| ||||||| ||||||| ||||||| |||||||
208 GATCCTGCTGCTGTGGATGGAGCCCTGGCAGCTGCTATGAAGTACATTT 159
||||| ||||||| ||||||| ||||||| ||||||| |||||||
31 eaGTYrPheTYrAsnArGThrSerLYsArGcYsGluThrPheValPhes 48
||||| ||||||| ||||||| ||||||| ||||||| |||||||
158 TAGATTTTTTCTACACCAACTGCCAACAGAGTGCAGATTTTCTGTTC 109
||||| ||||||| ||||||| ||||||| ||||||| |||||||
48 erGlyCysAsnGlyAsnLeuAsnPhelYsLeuLYsileGluArGlu 64
||||| ||||||| ||||||| ||||||| ||||||| |||||||
108 CTGGCTGCAATGCAACTTTTAATAACTTCAAGCTGAAATAGACTGTGAT 59
||||| ||||||| ||||||| ||||||| ||||||| |||||||
65 ValAlaCysValAlaLYsLYsProPro 74
||||| ||||||| ||||||| ||||||| ||||||| |||||||
58 GTAACCTTGTCACGAAACATACAAAAACCCCT 29
||||| ||||||| ||||||| ||||||| ||||||| |||||||

seq_name: gb_est15:AI015730
seq_documentation_block:
LOCUS AI015730 496 bp mRNA EST 27-AUG-1998
DEFINITION ot71b08.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:1622199 3' similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN
); mRNA sequence.
ACCESSION AI015730
VERSION AI015730.1 GI:3230066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 496)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2157 Std Error: 0.00
Seq primer: -40ml3 fwd ET from Amersham
High quality sequence stop: 272.
FEATURES
source
location/Qualifiers
1..496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1622199"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/lab_host="8-9 weeks"
/lab_host="DHI0B"
/notes="Vector: p7T73D-Pac (Pharmacia) with a modified
polylinker; Site:1: Not 1; Site:2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTCACCAATCTGAAGTGGAGCGCCCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 105 a 145 c 128 g 118 t
ORIGIN

alignment_scores:
Quality: 132.00 Length: 54
Ratio: 3.882 Gaps: 0
Percent Similarity: 62.963 Percent Identity: 48.148

```

```
alignment_block:
US-09-215-435-177_COPY_25_99 x AI015730/rev ..
Align seg 1/1 to reverse of: AI015730 from: 1 to: 496

14 LysAspProCysLysLeuAspMetAsnPheGlySerCysTyrGluValH1 30
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
493 GAAGATTCCTCGCAGCTGGGCTACTCGCGGGTCTCTGCATGGGATGAC 444
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

30 sPheArgTyrPheTyrAsnArgThrSerLysArgCysGluThrPheValP 47
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 CAGCAGGATTTCTATATGTTACATCCATGCGCTCTGAGACTTTCCAGT 394
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

47 heSerGlyCysAsnGlyAsnLeuAsnAsnPheLysLeuLysIleGluArg 63
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 ACGCGGCTGCATGGCGCAACGGTACAACTTCGTCACAGAAAGAGTGT 344
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

64 GluValAlaCys 67
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 CTCGACACCTGC 332

seq_name: gb_est18:AI256949

seq_documentation_block:
LOCUS AI256949 552 bp mRNA EST 17-NOV-1998
DEFINITION LP04037.5prime LP Drosophila melanogaster larval-early pupal pot2
Drosophila melanogaster cDNA clone LP04037 5prime, mRNA sequence.
ACCESSION AI256949
VERSION AI256949.1 GI:3864474
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 552)
AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C., Brokstein
, P., Lewis, S. and Rubin, G.M.
TITLE BDGP/HHMI Drosophila EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Sequence is the complete cDNA insert. hit genomic sequence DS00216;
hit genomic sequence DS02911; hit genomic sequence DS03364
plate: 40 row: D column: 1
High quality sequence stop: 552.
FEATURES
Location/Qualifiers
1..552
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LP04037"
/clone_lib="LP Drosophila melanogaster larval-early pupal
pot2"
/sex="male and female"
/dev_stage="larvaepupae"
/lab_host="DH5-alpha"
/note="Organ: whole body; Vector: pot2; Site:1: EcoRI;
Site:2: XhoI; Sized fractionated cDNAs were directly
ligated into pot2. Plasmid cDNA library."

BASE COUNT 154 a 124 c 122 g 152 t
ORIGIN

alignment_scores:
Quality: 131.00 Length: 56
Ratio: 3.639 Gaps: 0
Percent Similarity: 64.286 Percent Identity: 41.071

seq_documentation_block:
LOCUS D83828 552 bp mRNA EST 15-APR-1996
DEFINITION HUMHBC4837 Human pancreatic islet Homo sapiens cDNA similar to pp5,
mRNA sequence.
ACCESSION D83828
VERSION D83828.1 GI:1262515
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 552)
AUTHORS Takeda, J.
TITLE Human pancreatic islet ESTs
JOURNAL Unpublished (1995)
COMMENT Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@eb.gunma-u.ac.jp.
FEATURES
Location/Qualifiers
1..552
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human pancreatic islet"
/note="Vector: Lambda ZAPII; Site:1: Eco RI; Site:2: Xho
I; mRNA was prepared from normal adult human islets. cDNA
was directionally synthesized from the xho I in the vector
to the EcoRI site. cDNA was size fractionated to remove
sequences <1000 bp in size."

BASE COUNT 146 a 129 c 136 g 135 t
ORIGIN

alignment_scores:
Quality: 129.00 Length: 70
Ratio: 3.000 Gaps: 0
Percent Similarity: 61.429 Percent Identity: 34.286

alignment_block:
US-09-215-435-177_COPY_25_99 x D83828 ..
Align seg 1/1 to: D83828 from: 1 to: 552

6 AlaGluLysIleCysGlyAspLeuLysAspProCysLysLeuAspMetAs 22
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 GCTCAGGAGCCAAACAGGAAATAACGCGAGATCTGCTCTCCCTCCCTAAA 50
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
22 nPheGlySerCysTyrGluValH1sPheArgTyrPheTyrAsnArgThrS 39
```



```

||||| ||| ::| ::| ||||| ::| ||| ::| ||| ::| ||| ::| |||
51 CTACGACCCCTCGCGGCCTACTTCTCGTTACTACTACGACGGTACA 100

39 erLysArgCysGluThrPheValPheSerGlyCysAsnGlyAsnLeuAsn 55
||||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
101 CGCAGAGCTGGCGCAGTCTGTGTACGGGGGCTCGGAGGGCAACGCCAAC 150

56 AsnPheLysLeuLysIleGluArgGluValAlaLysCysValAlaLysTyrly 72
||||| ::| ::| |||||
151 AATTCTACACCTGGAGGCTTGCACACNATGCTTCTCGAGGATANAAAA 200

72 sProProArg 75
| ||| ::|
201 ACTTCCCAAA 210

seq_name: gb_est109:T74655

seq_documentation_block:
LOCUS T74655 350 bp mRNA EST 02-MAR-1995
DEFINITION yc57ell.r1 Stratagene liver (#937224) Homo sapiens cDNA clone
IMAGE:84812 5' similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN);,
mRNA sequence.
ACCESSION T74655
VERSION T74655.1 GI:691330
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 350)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevasakis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 62
High quality sequence stops: 281 Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL ; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 62 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 281.
Location/Qualifiers
1. .350
/organism="Homo sapiens"
/db_xref="GDB:501869"
/db_xref="taxon:9606"
/clone="IMAGE:84812"
/clone_lib="Stratagene liver (#937224)"
/sex="male"
/dev_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Hepatotomy from normal male caucasian. Average insert
size: 1.1 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5'
GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
CTCAGCTTTTCTTTTCTTTT 3'"
BASE COUNT 78 a 92 c 103 g 76 t 1 others
ORIGIN

```

```

Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909
alignment_block:
US-09-215-435-177_COPY_25_99 x T74655
Align seg 1/1 to: T74655 from: 1 to: 350

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysly 18
||||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
14 GTAACCTGAAGTCAACCAAGAAA.....GAAGATTCTCTGCCCA 48

18 sleuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 35
||||| ::| ::| ||| ::| ::| ::| ::| ::| ::| ::| ::|
49 GCTGGGCTACTCGCGCGTCCCTCGATGGGAATGACCAACGATATTCT 98

35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
||||| ||||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
99 ATAATGGTACATCCATGGCTGTGAGACTTTCAGTACGCGGCTGCATG 148

52 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaLys 67
||||| ||||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
149 GCACAACGGTAACTTCTGTCACAGAAAAGAGTCTCTGCAGACCTGC 196

seq_name: gb_est7:AA424642

seq_documentation_block:
LOCUS AA424642 413 bp mRNA EST 16-OCT-1997
DEFINITION z991d02.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:767139 5'
similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN);, mRNA sequence.
ACCESSION AA424642
VERSION AA424642.1 GI:2103595
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST project 1997
Unpublished (1997)
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham.
Location/Qualifiers
1. .413
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:767139"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,

```

alignment_scores:

```
Align seg 1/1 to: AF063548 from: 1 to: 511
2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 GTAAGTGAAGTCAACCAAGAAA.....GAAGATTCTCTGCCA 227

18 sLeuAspMetAsnPhcGlySerCysTyrGluValHisPheArgTyrPheT 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 GCTGGGCTACTCGCGCGTCCCTGCGATGGGAATGACACAGGATTTCT 277

35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 ATAATGGTACATCCATCGCTGTGAGACTTTCAGTACGGCGGCTGCATG 327

52 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 GGCAACGGTAACAACACTTCGTACAGAAAAGGAGTGTCTCCAGACCTGC 375

seq_name: gb_est16:A1110674

seq_documentation_block: 511 bp mRNA EST 11-NOV-1999
LOCUS A1110674 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
DEFINITION HA0104 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION A1110674
VERSION A1110674.1 GI:6359539
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
CONTACT: Yongtao Yu
DEPARTMENT: Department of Hematology
INSTITUTION: Beijing Institute of Radiation Medicine
ADDRESS: 27 Taiping Road, Beijing 100850, P.R.China
TEL: 0086-10-68159479
FAX: 0086-10-68214653
EMAIL: yyt48@yahoo.com.
FEATURES
source
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="vector: pCDNA1"
BASE COUNT 114 a 140 c 155 g 102 t
ORIGIN

alignment_scores:
Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x A1110674 ..
Align seg 1/1 to: A1110674 from: 1 to: 511

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 GTAAGTGAAGTCAACCAAGAAA.....GAAGATTCTCTGCCA 227

18 sLeuAspMetAsnPhcGlySerCysTyrGluValHisPheArgTyrPheT 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 GCTGGGCTACTCGCGCGTCCCTGCGATGGGAATGACACAGGATTTCT 277

35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 ATAATGGTACATCCATCGCTGTGAGACTTTCAGTACGGCGGCTGCATG 327

52 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 GGCAACGGTAACAACACTTCGTACAGAAAAGGAGTGTCTCCAGACCTGC 375

seq_name: gb_est14:AF063548

seq_documentation_block: 511 bp mRNA EST 16-JUN-1999
LOCUS AF063548 Homo sapiens library (Yu Y) Homo sapiens cDNA clone HA0104
DEFINITION AF063548 Homo sapiens library (Yu Y) Homo sapiens cDNA clone HA0104
, mRNA sequence.
ACCESSION AF063548
VERSION AF063548.1 GI:5080918
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS Yu, Y., Dong, C. and He, F.
TITLE Finding some ESTs
JOURNAL Unpublished (1998)
CONTACT: Yongtao Yu
DEPARTMENT: Department of Hematology
INSTITUTION: Beijing Institute of Radiation Medicine
ADDRESS: 27 Taiping Road, Beijing 100850, P.R.China
TEL: 0086-10-68159479
FAX: 0086-10-68214653
EMAIL: yyt48@yahoo.com.
FEATURES
source
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HA0104"
/clone_lib="Homo sapiens library (Yu Y)"
BASE COUNT 114 a 140 c 155 g 102 t
ORIGIN

alignment_scores:
Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x AF063548 ..
```

```
35 yrasnargthrserlyscysglutrphevalpsheserglycysasn 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
278 ATAATGGTACATCCATCGGCTGTGAGACTTTCAGTACGGCGGCTGCATG 327

52 GlyAsnLeuAsnAspPheLysLeuLysLeuGluArgGluValAlaCys 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 GGCAACGGTAACAACCTCGTCACAGAAAAGGAGTGCTGCAGACCTGC 375

seq_name: gb_est40:AV720273

seq_documentation_block:
LOCUS AV720273 520 bp mRNA EST 16-OCT-2000
DEFINITION AV720273 GLC Homo sapiens cDNA clone GLCGBA03 5', mRNA sequence.
ACCESSION AV720273
VERSION AV720273.1 GI:10817425
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 520)
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA GLC clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..520
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCGBA03"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 107 a 152 c 137 g 124 t
ORIGIN

alignment_scores:
Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x AV720273/rev ..
Align seg 1/1 to reverse of: AV720273 from: 1 to: 520

2 ValAsnLysIleAlaGluLysLeuCysGlyAspLeuLysAspProCysLy 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
512 GTAACCTGAAGTACCACAGAAA.....GAAGATTCCTGCCA 478

18 sLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 35
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 GCTGGGCTACTCGCGCGGTCCTGCATGGAAATGACACGACGATTTTCT 428

35 yrasnargthrserlyscysglutrphevalpsheserglycysasn 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 ATAATGGTACATCCATCGGCTGTGAGACTTTCAGTACGGCGGCTGCATG 478
```

```
52 GlyAsnLeuAsnAspPheLysLeuLysLeuGluArgGluValAlaCys 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
377 GGCAACGGTAACAACCTCGTCACAGAAAAGGAGTGCTGCAGACCTGC 330

seq_name: gb_est17:AI207711

seq_documentation_block:
LOCUS AI207711 521 bp mRNA EST 11-NOV-1999
DEFINITION HA3093 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION AI207711
VERSION AI207711.1 GI:6361726
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 521)
AUTHORS Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Contact: Chenggang Zhang
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Email: zhang.chenggang@hotmail.com.
FEATURES
Location/Qualifiers
source
1..521
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="vector: pCDNAL"

BASE COUNT 116 a 142 c 157 g 106 t
ORIGIN

alignment_scores:
Quality: 128.50 Length: 66
Ratio: 3.212 Gaps: 1
Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x AI207711 ..
Align seg 1/1 to: AI207711 from: 1 to: 521

2 ValAsnLysIleAlaGluLysLeuCysGlyAspLeuLysAspProCysLy 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 GTAACCTGAAGTACCACAGAAA.....GAAGATTCCTGCCA 227

18 sLeuAspMetAsnPheGlySerCysTyrGluValHisPheArgTyrPheT 35
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
228 GCTGGGCTACTCGCGCGGTCCTGCATGGAAATGACACGACGATTTTCT 277

35 yrasnargthrserlyscysglutrphevalpsheserglycysasn 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
278 ATAATGGTACATCCATCGGCTGTGAGACTTTCAGTACGGCGGCTGCATG 327

52 GlyAsnLeuAsnAspPheLysLeuLysLeuGluArgGluValAlaCys 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 GGCAACGGTAACAACCTCGTCACAGAAAAGGAGTGCTGCAGACCTGC 375

seq_name: gb_est17:AI207590

seq_documentation_block:
LOCUS AI207590 571 bp mRNA EST 11-NOV-1999
DEFINITION HA2954 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION AI207590
VERSION AI207590.1 GI:6361600
```

Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzgchgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source
 1.. 626
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GLCFQ208"
 /clone_lib="GLC"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 163 a 157 c 183 g 120 t 3 others
 ORIGIN

alignment_scores:
 Quality: 128.50 Length: 66
 Ratio: 3.212 Gaps: 1
 Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
 US-09-215-435-177_COPY_25_99 x AV658823 ..
 Align seg 1/1 to: AV658823 from: 1 to: 626

2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLys 18
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 98 GTAACCTGAAGTACCACGAAGA.....GAAGATTCTCTGCCA 132

18 sLeuAspMetAsnPhleGlySerCysTyrGluValHisPheArgTyrPheT 35
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 133 GCTGGGGTACTCGCCGGTCCCTGCATGGGAATGACACGAGGATTTTCT 182

35 yrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 183 ATAATGGTACATCCATGGCTTGTGAGACTTCCAGTACGGCGGTGCGATG 232

52 GlyAsnLeuAsnAsnPhelysLeuLysIleGluArgGluValAlaCys 67
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 233 GGCAACGGTACAACTTCGTGCACGAAGAAGGAGTCTCTGCAGACCTGC 280

seq_name: gb_est25:AI807643

seq_documentation_block:
 LOCUS AI807643 644 bp mRNA EST 19-DEC-1999
 DEFINITION wf49404 xl Soares.NFL.T.GBC.S1 Homo sapiens cDNA clone
 IMAGE:2358919 3' similar to gb:X04225 ALPHA-1-MICROGLOBULIN (HUMAN
);,, mRNA sequence.

ACCESSION AI807643
 VERSION AI807643.1 GI:5394209
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (Bases 1 to 644)
 NCICGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@email.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1328 Std Error: 0.00
 Seq primer: -400P from Gbco

```

FEATURES             High quality sequence stop: 462.
Location/Qualifiers
Source               1. .644
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2358919"
                    /clone_lib="Soares_NFL_T_GBC_S1"
                    /lab_host="DH10B"
                    /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                    a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
                    Equal amounts of plasmid DNA from three normalized
                    libraries (fetal lung NBHL19W, testis NHT, and B-cell
                    NCI_CCAP_GCB1) were mixed, and ss circles were made in
                    vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo."
BASE COUNT          128 a 195 c 173 g 148 t
ORIGIN

alignment_scores:
  Quality: 128.50      Length: 66
  Ratio: 3.212        Gaps: 1
  Percent Similarity: 60.606 Percent Identity: 40.909

alignment_block:
US-09-215-435-177_COPY_25_99 x AI807643/rev ..

Align seg 1/1 to reverse of: AI807643 from: 1 to: 644

      2 ValAsnLysIleAlaGluLysIleCysGlyAspLeuLysAspProCysLy 18
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      508 GTAACGTGAAGTCAACCAAGAAA.....GAAGATTCTCGCCA 474

      18 sLeuAspMetAsnPheGlySerCysTyrgluValHisPheArgTyrPheT 35
      :|||:   :|||:   |||:   :|||:   :|||:   |||:   |||:
      473 GCTGGGCTACTCGCGCGGTCCCTGCTGATGGGAATGACCAGCAGGTATTCT 424

      35 YrAsnArgThrSerLysArgCysGluThrPheValPheSerGlyCysAsn 51
      |||||: |||||: |||||: |||||: |||||: |||||: |||||:
      423 ATAATGGTACATCCATGGCCTCTGAGACTTTCAGTACGGCGGCTGCATG 374

      52 GlyAsnLeuAsnAsnPhelLysLeuLysIleGluArgGluValAlaCys 67
      |||||: |||||: |||||: |||||: |||||: |||||: |||||:
      373 GGCAACGGTAACTTCGTCACAGAAAAGAGAGTCTCTGCAGACCTGC 326

```

THIS PAGE BLANK (USPTO)


```

alignment_scores:      Length: 349
      Quality: 345.00      Gaps: 2
      Ratio: 3.881      Percent Identity: 26.074
Percent Similarity: 26.934

alignment_block:
US-09-215-435-179 x AL157823 ..

Align seg 1/1 to: AL157823 from: 1 to: 127384

22 ArgGlySerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSe 38
||| |||||::: :||| ||||| ||||| |||||
87854 CGGTGGAGCCTGGCGGTGCCGTGTCCTCCCTCAGGAGCTCAGAGATC 87903

38 rCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProA 55
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87904 TTGCATCCGNAACCAGGACTGCGAGACTGGCTGCTGCCAAGCTGCTCCAG 87953

55 spAsnCysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCys 71
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87954 ACAATTGCGATGTCACATGCGCGGAGAGGGGTCCGAGGGCACTCTGTGT 88003

72 GlnThrGlnVal..Phe..... 76
||||| ||||| ::|
88004 CAACGCAGGTGGGTATCGCCGCCCGGGGGAGCCAGAGGGATCCAGGG 88053

76 ..... 76
88054 GAAGTGGAGCCAGGCGCGGCCAGATTCTGGGGAGGAAGAGGGTAG 88103

76 ..... 76
88104 GTGGGGGAAATGCCCTGGAGCTTGAATTCCCCCTGGGACTGCACCTC 88153

76 ..... 76
88154 CAAGCCTTCTGGGGAGCCTGTGGCAGGAACCAATATGTGAGCCACCCT 88203

76 ..... 76
88204 AGGGTTTGGAACTCTGGATTGGGGGGGAAGCGGATGTTTCACTTTCTA 88253

76 ..... 76
88254 ATGTGGGTTTCACTTCTCCAGGAGCTTCTCAGCAAAATTCGGAAGTAGT 88303

76 ..... 76
88304 ACAGTAATGGGCACAGCTAGGTCCAGCCTCTGCCCCAGAGCCCGCTTT 88353

76 ..... 76
88354 CATCCAGCTCAGGAGTTCTGCTGTGGCCCCGGGTGTTCTCGTGGAGGAT 88403

76 ..... 76
88404 GGGCAGGCGACCTCTCAGCTGGGTGTGTAGGGAGGGAGTTGATTTGC 88453

76 ..... 76
88454 CCATGGAAGATTTAAAGACAAGACATAGAGTCTCGCATGTTGTGTGCC 88503

76 ..... 76
88504 ATGGACTTAGNACAGCAGATGGGTGAGTGACAGCAGGGTTTACAAGTTC 88553

76 ..... 76
88554 TGGAGCTGGGAGGACACTGAACAATCCAGGACCCACTGGAGAGGCAG 88603

76 ..... 76
88604 GCGAGCGCGCTGTCAGCAGCTGCGATATCTATCTATCTCCCTGGATC 88653

```

seq_name	gb_hgt22	AL512371
seq_documentation_block:		
LOCUS	AL512371	159482 bp DNA HTG 21-MAR-2001
DEFINITION	Homo sapiens chromosome 6 clone RP11-238F24, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.	
ACCESSION	AL512371	
VERSION	AL512371.4	GI:13274864
KEYWORDS	HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 159482)	
TITLE	Clark.G.	
JOURNAL	Direct Submission	
COMMENT	Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk On request, clonerquest@sanger.ac.uk On Mar 12, 2001 this sequence version replaced gi:12964529. ----- Genome Center Center: Sanger Centre Center code: SC Web site: http://www.sanger.ac.uk Contact: humquery@sanger.ac.uk ----- Project Information Center: project name: BA238F24 ----- Summary Statistics Assembly program: XGAP4; version 4.5 Sequencing vector: plasmid; L08752; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Consensus quality: 154970 bases at least Q40 Consensus quality: 156246 bases at least Q30 Consensus quality: 156738 bases at least Q20 Insert size: 159282; sum-of-contigs Insert size: 165487; 7.5% error; agarose-fp Quality coverage: 6.35x in Q20 bases; sum-of-contigs Quality coverage: 6.11x in Q20 bases; agarose-fp ----- * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. * 3067 3166: gap of 100 bp * 3167 155002: contig of 151836 bp in length * 155003 155102: gap of 100 bp * 155103 159482: contig of 4380 bp in length.	


```

fragment_chain:2"
76344..81840
/note="assembly_fragment:00425
fragment_chain:2"
81941..86241
/note="assembly_fragment:02764
fragment_chain:2"
86342..129124
/note="assembly_fragment:00846
fragment_chain:2"
129225..131985
/note="assembly_fragment:01963"
132086..151439
/note="assembly_fragment:02120
fragment_chain:3"
151540..168884
/note="assembly_fragment:01758
fragment_chain:3
clone_end:17
vector_side:right"
BASE COUNT 46011 a 38577 c 39446 g 43544 t 1306 others
ORIGIN

alignment_scores:
    Quality: 346.00      Length: 349
    Ratio: 3.681        Caps: 2
Percent Similarity: 26.934  Percent Identity: 26.074

alignment_block:
US-09-215-435-179 x AL590400 ..

Align seg 1/1 to: AL590400 from: 1 to: 168884

22  ArGlySerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSe 38
|||||:|||||:
114892  CGGTGGAGCTGGCGGTGCGGTGCTCCTCCCTCCAGAGACTCAAGAGTCT 114941

38  rCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProA 55
|||||:|||||:|||||:|||||:|||||:|||||:
114942  TTGCATCCGGACACAGACTTCGAGACTGGCTGCTGCCAACGGTGTCCAG 114991

55  sPAsnCysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCys 71
|||||:|||||:|||||:|||||:|||||:|||||:
114992  ACAATTGCAGTCCGACTCGCGGAGAAAGGGTCCGAGGCACTGTGTGT 115041

72  GlnThrGlnVal...Phe..... 76
|||||:|||||:|||||:|||||:|||||:|||||:
115042  CAAACGACAGTGGGTATGCGCCCGGGGGGAGCCAGAGGGATCCAGGG 115091

76 ..... 76

115092  GAAGTGGAGCCAGGAGGCGCCAGATTCTCTGGGAGGAAGAAGGGTAG 115141

76 ..... 76

115142  GTGGCGGAAATGCCCTGGAGCCTGGAAATTCGCCCGTGGGACTCCACCTC 115191

76 ..... 76

115192  CAAGCCTTCTGGGGAGCCTGTGGACGGAACCAATATGTTAGCCACCTC 115241

76 ..... 76

115242  AGGGTTTTGGAATCCTGGATTTGGCGGGGAGCGGATGTTTTCACCTTTCTA 115291

76 ..... 76

115292  ATGTGGGTTCACATTTCTCCAGGAGTCTTCTCAGCAAATTCGGAAGTAGT 115341

76 ..... 76

115342  ACAGTAAATGGGCACAGCTAGGTCAGGCTCTTGCCCGAGAGCCCGTCTT 115391

```

```
76 ..... 76
115392 CATCCAGCTCAGGAGTTCTGCTGTGSCCCGGGTGTTCTCGTGGAGGAT 115441
76 ..... 76
115442 GGCAGAGCAGCCTCTGAGCTGGGTGTAGGGAGGAGTTGATTGC 115491
76 ..... 76
115492 CCATGAAGATTAAAGACACACATAGAGTCTCGCATGGTGTGTCC 115541
76 ..... 76
115542 ATGGACTTAGCAGCAGCAGATGGTCAAGTACAGGCGAGGTTACAGGTC 115591
76 ..... 76
115592 TGGAGCTGGGAGCAGACTGAACAATCCAGGACCGCCTGGAGAGGCAG 115641
76 ..... 76
115642 CCAGAGTGCCTCTGAGCAGCTGGCAAGCTATCATTTATCCTCGATC 115691
76 ..... 76
115692 CTGGCTGTGGCAAGCAGATGGCCCTGGAGCTGGGCTGGGCTGCTTA 115741
76 ..... 76
115742 GGGGCTCAGGAAGAGCTGCGCCAGGTCGAGTCAAAAGTCTGCTTT 115791
77 ..... PheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeu 89
115792 CCCAGTGTTCCTTTGGCCAAATATAGAGCGTGTCCCTGCTGCGGAACCTG 115841
90 ThrCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyAr 106
115842 ACTGTATATATCAAGAATGAGAAATGGCTAGCATCGCCTATGCGCG 115891
106 gCysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 121
115892 TTGTCAAAAATTTGGAAGCAGAAGTTGGCTAAGAAAATGTTCTTC 115937
seq_name: gb_om:CCCOLIP

seq_documentation_block:
LOCUS CCCOLIP 452 bp mRNA MAM 31-MAR-1995
DEFINITION Canine mRNA for colipase.
ACCESSION X53564
VERSION X53564.1 GI:841
KEYWORDS colipase.
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 452)
Direct Submission
Submitted (19-JUN-1990) Fukuoka S.-I., Scheele G.A., The Charles A.
Dana Research Institute, Harvard Medical School, Cell and Molecular
Biology, Dana 533, Beth Israel Hospital, 330 Brookline Ave.,
Boston, MA 02215, USA
2 (bases 1 to 452)
Fukuoka,S., Taniguchi,Y., Kitagawa,Y. and Scheele,G.
Full length cDNA sequence encoding canine pancreatic colipase
Nucleic Acids Res. 18 (18), 5549 (1990)
91016846
FEATURES
Location/Qualifiers
1..452
/organism="Canis familiaris"
/db_xref="taxon:9615"

/dev_stage="adult"
/tissue_type="pancreas"
/clone_lib="pUC-9"
/clone="pCL1"
1..51
/product="signal peptide"
1..339
/codon_start=1
/product="precursor peptide"
/protein_id="CAA37636.1"
/db_xref="GI:842"
/db_xref="SWISS-PROT:P19090"
/translation="MEKILVLLVALAVYAVDPDRGIIHLEDEGLCLNSVQCKSK
CHRATGLSLARCAPKASENSESAKTLGVYKPCERGLTCEGDKSIVGSIITNTNFG
VCHDAGRSKK"
52..336
/product="colipase"
429..435
/feature="polyA signal"
452
/feature="polyA site"
/feature="polyA site"
90 a 146 c 117 g 99 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 147.00 Length: 84
Ratio: 2.625 Gaps: 1
Percent Similarity: 66.667 Percent Identity: 36.905

alignment_block:
US-09-215-435-179 x CCCOLIP ..
Align seg 1/1 to: CCCOLIP from: 1 to: 452
33 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyC 49
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
82 CTGGAAGACGGCGAGCTGCTGCTGAACAGGCTCCAGTGAAGCAAGTG 131
49 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGly 66
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 CTGCCACCGCGCCACCGCGGCTGAGCCTGCGCGCTGCGCACCAAGCCA 181
66 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 82
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 GCAGAACACGAGTCTCTGCAACAGACGCTCTATGGGTTTACTACAAG 231
83 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsn.....GI 97
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 TGTCCCTGTGAGCGGGGCGCTGACCTGTGAGGGTGACAAGTCCATCGTGG 281
97 uLysTrpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnL 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
282 CTCCATCACAATACCAACTTTGGTGTCTGCCACGATGCTGGCGGCTCCA 331
114 vs 114
||
332 AG 333
seq_name: gb_om:RABCOLIPAS
seq_documentation_block:
LOCUS RABCOLIPAS 459 bp mRNA MAM 05-OCT-1993
DEFINITION Oryctolagus cuniculus pancreatic colipase mRNA, complete cds.
ACCESSION L06329
VERSION L06329.1 GI:164893
KEYWORDS colipase.
SOURCE Oryctolagus cuniculus male adult pancreas cDNA to mRNA.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
1 (bases 1 to 459)
REFERENCE
1 Colwell,N.S., Aleman-Gomez,J.A., Sasser,T. and Kumar,V.B.
```


REFERENCE 1 (bases 1 to 656)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences from normal pancreas tissue
JOURNAL Patent: WO 954446-A 16 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
source
1. .656
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 165 a 200 c 162 g 129 t
ORIGIN

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:
US-09-215-435-179 x AX014220 ..

Align seg 1/1 to: AX014220 from: 1 to: 656

12 LeuLeuPhePheLeuLeuPheLeuLeuThrArgGlySerLeuSerProth 28
98 ATCTGATCTCTGCTGTGTCGCCCTCTGTGGCTATGCAGCTCTCTGG 147
28 rLysTyrAsnLeu.....LeuGluleuLysGluSerCysIleArgA 42
148 CCCCAGGGGATCATTTATCAACCTGCAGAACGGTGAGCTGCATCAATA 197
42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
198 GTGCCAGGTGTAAGACGATTTGTCGACGATTCAGTGCCTGGCCCTG 247
59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
248 GCCCGTGCATCCATGCCAGGAGAGACGAGTGCTGTCTGTCAGAC 297
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
298 GCCTATGGGATTTACTACAAGTGTCCTGTGAGCGTGCCCTGACCTGT. 346
92 letYrSerLysAsnGluLysTyrLeu.....SerIleAla 103
347GAGGAGACAAGACCATCGTGGCTCCATCACCAACCAAC 388
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
389 TTGCGCATCTGCCATGACGCTGGACGCTCCAAG 421

seq_name: gb_rol:MCCOLIPAS

seq_documentation_block:
LOCUS MCCOLIPAS 490 bp mRNA ROD 21-NOV-1996
DEFINITION M.coypus mRNA for colipase.
ACCESSION X82998
VERSION X82998.1 GI:599866
KEYWORDS colipase.
SOURCE nutria.
ORGANISM Myocastor coypus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystricognathi; Myocastoridae;
Myocastor.
1 (bases 1 to 490)
REFERENCE 1 (bases 1 to 490)
AUTHORS Thirstrup,K., Carriere,F., Hjorth,S.A., Rasmussen,P.B.,
Nielsen,P.F., Ladefoged,C., Thim,L. and Boel,E.
TITLE Cloning and expression in insect cells of two pancreatic lipases
and a procolipase from Myocastor coypus
JOURNAL Eur. J. Biochem. 227 (1-2), 186-193 (1995)

MEDLINE 95154288
REFERENCE 2 (bases 1 to 490)
AUTHORS Thirstrup,K.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-1994) K. Thirstrup, Univ. Hospital, Lab. for
Molecular Endocrinology, Rigshospitalet 6321, Blegdamsvej 9,
DK-2100 Copenhagen, DENMARK
FEATURES
source
1. .490
/organism="Myocastor coypus"
/db_xref="taxon:10157"
/tissue_type="pancreas"
46. .381
/codon_start=1
/product="colipase"
/protein_id="CAA58119.1"
/db_xref="GI:599867"
/db_xref="SWISS-PROT:P42889"
/translation="MEKYLALVLTAVAYAAPDPRLIINLNGELCLNSAQKSOQ
QHDSPLGLARCADKARENSGSPQTYIYVLCPCERGLTCDGDKSIIGATNTNYG
ICQDPOSKK"
46. .96
97. .378
/note="putative propeptide: 97. .111"
/product="colipase"
97. .111
/note="putative propeptide"
/product="colipase"
BASE COUNT 96 a 181 c 120 g 93 t
ORIGIN

alignment_scores:
Quality: 134.50 Length: 95
Ratio: 2.360 Gaps: 2
Percent Similarity: 60.000 Percent Identity: 34.737

alignment_block:
US-09-215-435-179 x MCCOLIPAS ..

Align seg 1/1 to: MCCOLIPAS from: 1 to: 490

33 LeuGluleuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyC 49
127 CTGACAAACGGCGAGCTCTGCCTGAACAGCGCCAGTGCAGAGTCAGT 176
49 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 66
177 CTGCCAGCAGCAGACGCCCTCGGCCCTGGCCGCTGCGCGACAAGGCCA 226
66 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 82
227 GAGAGAACAGCGCTGCTCCCCACAGACGATCTATGGAATCTACTACT 276
83 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsn.....GI 97
277 TGTCCCTGCGAGCGGCCCTGACCTGTGACGGGGGACAAAGTCCATCGG 326
97 uLysTrpLeuSerIleAlaTyrGlyArgCysGln.....L 109
327 CGCCATCACCAACCAACCTACGCGATCTGCCAGGATCCTCAGTCCAAGA 376
109 ysIleGlyArgGlnLysLeuAlaLysLysMetPhe 120
377 AGTGAGGCCACACGCGCTGGCCCTGTCTCCCTTTC 411

seq_name: gb_om:ECCOLIPB

seq_documentation_block:
LOCUS ECCOLIPB 497 bp mRNA MAM 11-NOV-1994
DEFINITION E.caballus mRNA for colipase B.
ACCESSION X74344
VERSION X74344.1 GI:572678

Thu Nov 15 10:51:59 2001

```

KEYWORDS      colipase.
SOURCE        Equus caballus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE     Kerfelec.B.
AUTHORS       Direct Submission
JOURNAL       Submitted (28-JUL-1993) B. Kerfelec, CNRS, CDM, 31 Chemin Joseph
              Aguiier, 13402 Marseille, Cedex 9, FRANCE

REFERENCE     2 (bases 1 to 497)
AUTHORS       Crenon, I., Granon, S., Chapus, C. and Kerfelec, B.
JOURNAL       Molecular cloning and expression of two horse pancreatic cDNA
MEDLINE       encoding colipase A and B
FEATURES      Biochim. Biophys. Acta 1213 (3), 357-360 (1994)
              94325330
              Location/Qualifiers
                source
                  1..497
                  /organism="Equus caballus"
                  /db_xref="taxon:9796"
                  /tissue_type="pancreas"
                  <1..327
                  /codon_start=1
                  /product="colipase B"
                  /protein_id="CAA52391.1"
                  /db_xref="GI:572679"
                  /db_xref="SWISS-PROT:P02705"
                  /translation="LALLLVALLVAVAPDPGRGIIINLDEGEICMNSAQCKSECCHRE
                  SLSLARCAAKASENSECSAFTLVGVYKPCERGLTCQVDKTLVGSIMNTNFGICPD
                  AARSEE"
                CDS
                  118 a 161 c 125 g 93 t
                  BASE COUNT 118 a 161 c 125 g 93 t
                  ORIGIN
                    alignment_scores:
                      Quality: 134.00      Length: 63
                      Ratio: 3.045          Gaps: 0
                      Percent Similarity: 69.841 Percent Identity: 39.683

                    alignment_block:
                      US-09-215-435-179 x ECCOLIPB ..
                      Align seg 1/1 to: ECCOLIPB from: 1 to: 497
                        33 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 49
                        ||||| ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        70 TTGGAGCGCGGCGAGATCTGCTGATGAACAGCGCCAGTGCAGAGCGGAATG 119
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        49 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 66
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        120 CTGCCACCGGGAAGAGCGCTGAGCGTGGCCGCTGCGCAGCAAGCA 169
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        66 erGluGlySerLeuCysGlnThrGlnValPhePheGlyClnTyrArgAla 82
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        170 GCGAAGACAGCGAGTCTGCTGCTGACACTCTACGGGGTTACTACAA 219
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        83 CysProCysLeuArgAsnLeuThrCysIleTyrSerLys 95
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        220 TGTCCCTGTGAGCGGGCGCTGACCTGCCAGTGGACAA 258
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

                    seq_name: gb_om:AF148567

                    seq_documentation_block:
                      LOCUS AF148567 510 bp mRNA MAM 06-MAY-2000
                      DEFINITION Sus scrofa pancreatic colipase mRNA, complete cds.
                      ACCESSION AF148567
                      VERSION AF148567.1 GI:7711135
                      KEYWORDS pig.
                      SOURCE Sus scrofa
                      ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
                      REFERENCE 1 (bases 1 to 510)

AUTHORS       Darnis, S. and Chaix, J.C.
TITLE         Cloning, sequencing and functional expression of porcine pancreatic
              colipase
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 510)
AUTHORS       Darnis, S. and Chaix, J.C.
JOURNAL       Direct Submission
TITLE         Submitted (04-MAY-1999) LBBN, CNRS, ESA6033, Av Escadrille Normandie
              Niemen, Marseille 13397, France
FEATURES      Location/Qualifiers
                source
                  1..510
                  /organism="Sus scrofa"
                  /db_xref="taxon:9823"
                  /tissue_type="pancreas"
                  44..382
                  /note="cofactor; triglyceride lipase"
                  /codon_start=1
                  /product="colipase"
                  /protein_id="AAF67823.1"
                  /db_xref="GI:7711136"
                  /translation="MEKVLALLVLTIVAVAPDPGRGIIINLDEGEICLNSAQCKSWC
                  COHDYLSRSRALKARENSECSAFTLVGVYKPCERGLTCEGDKSLVGSITNTNFG
                  ICHDVGRSSD"
                CDS
                  117 a 160 c 121 g 112 t
                  BASE COUNT 117 a 160 c 121 g 112 t
                  ORIGIN
                    alignment_scores:
                      Quality: 133.00      Length: 85
                      Ratio: 2.463          Gaps: 2
                      Percent Similarity: 63.529 Percent Identity: 34.118

                    alignment_block:
                      US-09-215-435-179 x AF148567 ..
                      Align seg 1/1 to: AF148567 from: 1 to: 510
                        33 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 49
                        ||||| ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        125 CTGGATGAGCGCGAGCTGCTGCTGAACAGTCCCGAGTGCAGACGCAACTG 174
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        49 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 66
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        175 CTGCCAGCATGACACAACTCTGCGCTGCGCAGTCAAGGCA 224
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        66 erGluGlySerLeuCysGlnThrGlnValPhePheGlyClnTyrArgAla 82
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        225 GAGAGAAGACGCGAGTGTCTGCTTCACGCTCTATGGGGTTTACTACAA 274
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        83 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 99
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        275 TGTCCCTGTGACCGGGCGCTGACCTGT.....GAGGGGCAAGAG 315
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        99 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 111
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        316 TCTCGTGGGCTCCATCCCAACCACTTGTGATATCTGCCATGATGTG 365
                        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                        111 lyArg 112
                        |||||
                        366 GACGC 370

                    seq_name: gb_om:ECCOLIPA

                    seq_documentation_block:
                      LOCUS ECCOLIPA 472 bp mRNA MAM 28-JUL-1995
                      DEFINITION E. caballus mRNA for colipase A.
                      ACCESSION X74503
                      VERSION X74503.1 GI:572676
                      KEYWORDS anchoring protein; colipase A; secreted protein; trlacylglycerol
                      digestion.
                      SOURCE horse.
                      ORGANISM Equus caballus
                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

seq_documentation_block:	LOCUS	RATCOLQ	492 bp	mRNA	ROD	04-AUG-1994
	DEFINITION	Rat colipase mRNA, complete cds.				
	ACCESSION	M58370				
	VERSION	M58370.1 GI:203504				
	KEYWORDS	colipase.				
	SOURCE	Rat, cDNA to mRNA.				
	ORGANISM	Rattus norvegicus				
	REFERENCE	1 (bases 1 to 492)				
	AUTHORS	Payne,R.M., Sims,H.F., Jennens,M.L. and Lowe,M.E.				
	TITLE	Rat pancreatic lipase and two related proteins: enzymatic properties and mRNA expression during development				
	JOURNAL	Am. J. Physiol. 266 (5 Pt 1), G914-G921 (1994)				
	MEDLINE	94262798				
	FEATURES	Location/Qualifiers				
	source	1..492				
		/organism="Rattus norvegicus"				
		/db_xref="taxon:10116"				
	CDS	59..397				
		/codon_start=1				
		/product="colipase"				
		/protein_id="AAA20505.1"				
		/db_xref="gi:203505"				
		/translation="MKVLVLLVTLVAVAYAFGRGLFINLEDGEICVNSMOCKSRG COHDTILGIARCTHKAMENSECPKLYIVYPCBRLTCEGDRSLIIGAITNTNG VCLDSTRSKQ"				
	BASE COUNT	107 a 163 c 119 g 103 t				
	ORIGIN					
	alignment_scores:					
	Quality:	122.50	Length:	122		
	Ratio:	1.828	Gaps:	3		
	Percent Similarity:	54.918	Percent Identity:	27.869		
	alignment_block:					
	US-09-215-435-179 x RATCOLQ	..				
	Align seg 1/1 to: RATCOLQ	from: 1 to: 492				
	3	LeuProGlnTrpLeuLeuLeuLeuPheLeuLeuPhePheLeuPheLeuPheLeu	19			
	47	GTACCACTACCATCAAGGTCCTGTTGTTCTGCTTAACCTCGTTGC	96			
	19	uLeuThrArgGlySerLeuSerProThrLysTyrAsnLeu...LeuGluL	35			
	97	CGTGGCCTATGCAGCTCCTGGACCCCGGGCTCTTTTATCAACCTGGAGG	146			
	35	eulysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGln	51			
	147	ACGCTGAGATCTCGTAAACAGTATGCAGTGTAAAGACAGATGCTGCCAA	196			
	52	ArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerGluG1	68			
	197	CATGACACCATCTCGGCATCGCCCGGTGCACACACAGCCATGGAGAA	246			
	68	ySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysProC	85			
	247	CAGCGAGTGTCCCAAGACCCCTCTATGGGATCTACTACAGGTGCCCT	296			
	85	ysLeuArgAsnLeuThrCys.....Ile	92			
	297	GTAGCGGGCCCTGACCTGTGAGGGGACAGGACATCATTTGGCGCATC	346			
	93	TyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgCysGlnLy	109			
	347	ACCAACACCAAC.....TACGGGCTGTGCTCGCA	375			
	109	sileGlyArgGlnLys	114			
	seq_name:	gb_ro2:RATCOLQ				

```

376 CTCACCCCGCTCCAAG 391
seq_name: gb_ro2:RATCOLIP
seq_documentation_block: 446 bp mRNA ROD 27-APR-1993
LOCUS RATCOLIP
DEFINITION Rat colipase mRNA, complete cds.
ACCESSION M33333
VERSION M33333.1 GI:203502
KEYWORDS colipase.
SOURCE Rattus norvegicus (strain Wistar) (library: puc) pancreas cDNA to
        mRNA.
ORGANISM Rattus norvegicus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
        Rattus.
REFERENCE 1 (bases 1 to 446)
AUTHORS Wicker,C. and Puigserver,A.
TITLE Rat pancreatic colipase mRNA: Nucleotide sequence of a cDNA clone
        and nutritional regulation by a lipidic diet
JOURNAL Biochem. Biophys. Res. Commun. 167, 130-136 (1990)
MEDLINE 90179738
FEATURES
        Location/Qualifiers
            1..446
            /organism="Rattus norvegicus"
            /strain="Wistar"
            /db_xref="taxon:10116"
            /tissue_type="pancreas"
            /tissue_lib="puc"
            <1..446
            /note="colipase mRNA"
            18..356
            /codon_start=1
            /product="colipase"
            /protein_id="AAA40943.1"
            /db_xref="GI:203503"
            /translation="MKVLVLLVAVAYAVPGPGLFINLEDGEICVNSMOCKSRG
            CQHDITLGIARTCHKAMENSECSPTLGIYRCPGRGLCEGDRSIGAITNTNG
            VLDSTRSQ"
            sig_peptide
            mat_peptide 69..353
            /product="colipase"
            polyA_signal 424..430
            polyA_site 446
            BASE COUNT 96 a 141 c 113 g 96 t
            ORIGIN

alignment_scores:
        Quality: 122.00 Length: 91
        Ratio: 2.392 Gaps: 2
        Percent Similarity: 56.044 Percent Identity: 31.868

alignment_block:
US-09-215-435-179 x RATCOLIP ..
Align seg 1/1 to: RATCOLIP from: 1 to: 446
33 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCY 49
||||| ||| |||||:|||||:|||||:|||||:|||||:
99 CTGGAGCGGTGAGATCTGCGTAAACAGTATGCAGTGTGAAGGACGATG 148
|||||:|||||:|||||:|||||:|||||:|||||:
49 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 66
|||||:|||||:|||||:|||||:|||||:|||||:
149 CTGCCAACATGACACCATCTCTGGGCATCGCCCGGTGCACACAAAGGCA 198
|||||:|||||:|||||:|||||:|||||:|||||:
66 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnThrArgAla 82
|||||:|||||:|||||:|||||:|||||:|||||:
199 TGGAGAACACGAGTGTGTCGCCAAACACCCCTCTATGGGATCTACTACAGG 248
|||||:|||||:|||||:|||||:|||||:|||||:
83 CysProCysLeuArgAsnLeuThrCys.....
|||||:|||||:|||||:|||||:|||||:
249 TGTCCCTGTGAGGGGCGCTGACCTGTGAGGGGACAGGAGCATCATTTGG 298

```

```

92 .....IleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 107
||| :||| |||||
299 CGCATCAACCAACACCAAC.....TACGGCGTCT 327

107 ysglnLysIleGlyArgGlnLys 114
|| :||| |||||
328 GCCTCGACTCCACCGCTCCAG 350

seq_name: gb_in3:CEW02A2
seq_documentation_block:
LOCUS CEW02A2 35879 bp DNA INV 27-OCT-2000
DEFINITION Caenorhabditis elegans cosmid W02A2, complete sequence.
ACCESSION Z82286
VERSION Z82286.1 GI:1914563
KEYWORDS HTG.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
        Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida;
        Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 35879)
AUTHORS none.
TITLE Genome sequence of the nematode C. elegans: a platform for
        investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99069613
REMARK The C. elegans Sequencing Consortium.
        Erratum: [[published errata appear in Science 1999 Jan
        1:283(5398):35 and 1999 Mar 26:283(5410):2103 and 1999 Sep
        3:285(5433):1493]]
REFERENCE 2 (bases 1 to 35879)
AUTHORS Ainscough,R.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1996) Nematode Sequencing Project, Sanger Centre,
        Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
        Washington University, St. Louis, MO 63110, USA. E-mail:
        jes@sanger.ac.uk or rw@nematode.wustl.edu
COMMENT On Mar 30, 1997 this sequence version replaced gi:1666624.
        Coding sequences below are predicted from computer analysis, using
        predictions from Genefinder (P. Green, U. Washington), and other
        available information.
        Current sequence finishing criteria for the C. elegans genome
        sequencing consortium are that all bases are either sequenced
        unambiguously on both strands, or on a single strand with both a
        dye primer and dye terminator reaction, from distinct subclones.
        Exceptions are indicated by an explicit note.
        This sequence is the entire insert of clone W02A2. The start of
        this sequence (1..106) overlaps with the end of sequence AL132901.
        The end of this sequence (35780..35879) overlaps with the start of
        sequence AL132865.
        For a graphical representation of this sequence and its analysis
        see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
        name=W02A2.
FEATURES
        Location/Qualifiers
            1..35879
            /organism="Caenorhabditis elegans"
            /db_xref="taxon:6239"
            /chromosome="IV"
            /clone="W02A2"
            join(830..1480,2149..2274,2875..3228)
            /gene="W02A2.1"
            /note="fat-2"
            join(830..1480,2149..2274,2875..3228)
            /gene="W02A2.1"
            /note="predicted using Genefinder
            contains similarity to pfam domain: PF00487 (Fatty acid
            desaturase), Score=108.7, E-value=5.7e-32, N=1
            cDNA EST EMBL:Z14917 comes from this gene
            cDNA EST EMBL:M75805 comes from this gene
            cDNA EST EMBL:M89208 comes from this gene
            cDNA EST EMBL:M89244 comes from this gene
            cDNA EST yk13hl.5 comes from this gene

```


Thu Nov 15 10:51:59 2001

```

US-09-215-435-179 x CEW02A2/rev ..
Align seg 1/1 to reverse of: CEW02A2 from: 1 to: 35879

6 TrpLeuLeuLeuPheLeuPhePhePheLeuLeuLeuThrAr 22
  ::::: ||||| ::::: ||||| ::::: |||||
7871 TACCTTTCACTCTTCAACCTTTTGTATCTTCTTTATAGGAACCTT 7822

22 g.GlySerLeuSerProThrLysTyrAsn..... 31
  ||| ::::: ||||| ::::: |||||
7821 GCGAGTCTGCCAATGCCAGGATACAACTTCTGCCCAACTGCCCTCGGCC 7772

32 .....LeuLeuLeuLysGluSerCysIleArgAsnGlnAspCy 45
  ||| ::::: ||||| ::::: |||||
7771 ATGCAATTCCTGCTGAGAGACCCGTGAGCTCCAGACAGTGCCTG 7722

45 sGluThrGly.....CysCysGlnArgAlaProAspAsnCysGlu 59
  ||| ::::: ||||| ::::: |||||
7721 TCCAAATGGAGACAAACTGTGTCTGCCCATGCCCGGAAATCCTCGAGC 7672

59 erHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVal 75
  ||| ::::: ||||| ::::: |||||
7671 CAATCTCTGCTCAAGAGTCCGAGATCTGCCAGCAACGACCATG 7622

76 PhePheGlyGlnTyrArgAlaCysProCysLeuArgAsn.....Le 89
  ||||| ::::: ||||| ::::: |||||
7621 AGATTCTCCCAATACCATCATGCCATGCCACAGAACCAACGACCTG 7572

89 uThrCysIleTyrSer.....LysAsnGluLysTrpLeuSerIleAla 104
  ||||| ::::: ||||| ::::: |||||
7571 CACATGCTCTACCAACATCAGCAGAAATGAGAGATCCTT...GTCTCCA 7525

104 yrGlyArgCysGlnLys 109
  ::::: ||||| ::::: |||||
7524 TCTCTGCTGCCAACAA 7508

seq_name: gb_htg24:CEY62E10

```

```

/clone-"Y62E10"
BASE COUNT 77287 a 44462 c 42975 g 79051 t 10442 others
ORIGIN

alignment_scores:
  Quality: 96.50      Length: 123
  Ratio: 1.485       Gaps: 6
  Percent Smilarity: 52.846   Percent Identity: 31.707

alignment_block:
US-09-215-435-179 x CEY62E10/rev ..
Align seg 1/1 to reverse of: CEY62E10 from: 1 to: 254217

6 TrpLeuLeuLeuPheLeuPhePhePheLeuLeuLeuThrAr 22
  ::::: ||||| ::::: ||||| ::::: |||||
2063 TACCTTTCACTCTTCAACCTTTTGTATCTTCTTTATAGGAACCTT 2014

22 g.GlySerLeuSerProThrLysTyrAsn..... 31
  ||| ::::: ||||| ::::: |||||
2013 GCGAGTCTGCCAATGCCAGGATACAACTTCTGCCCAACTGCCCTCGGCC 1964

32 .....LeuLeuGluLeuLysGluSerCysIleArgAsnGlnAspCy 45
  ||||| ::::: ||||| ::::: |||||
1963 ATGCAATTCCTCTGCTGAGAGACCCGTGAGCTCCAGACAGTGCCTG 1914

45 sGluThrGly.....CysCysGlnArgAlaProAspAsnCysGlu 59
  ||| ::::: ||||| ::::: |||||
1913 TCCAAATGGAGACAAACTGTGTCTGCCCATGCCCGGAAATCCTCGAGC 1864

59 erHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVal 75
  ||| ::::: ||||| ::::: |||||
1863 CAATCTCTGCTCAAGAGTCCGAGATCTGCCAGCAACGACCATG 1814

76 PhePheGlyGlnTyrArgAlaCysProCysLeuArgAsn.....Le 89
  ||||| ::::: ||||| ::::: |||||
1813 AGATTCTCCCAATACCATCATGCCATGCCACAGAACCAACGACCTG 1764

89 uThrCysIleTyrSer.....LysAsnGluLysTrpLeuSerIleAla 104
  ||||| ::::: ||||| ::::: |||||
1763 CACATGCTCTACCAACATCAGCAGAAATGAGAGATCCTT...GTCTCCA 1717

104 yrGlyArgCysGlnLys 109
  ::::: ||||| ::::: |||||
1716 TCTCTGCTGCCAACAA 1700

```

```

seq_documentation_block:
LOCUS CEY62E10 254217 bp DNA HTG 06-SEP-1999
DEFINITION Caenorhabditis elegans chromosome IV clone Y62E10, *** SEQUENCING
IN PROGRESS ***, in unordered pieces.

ACCESSION AL031580
VERSION AL031580.13 GI:5823679
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1. (bases 1 to 254217)
Lloyd,C.
Direct Submission
Submitted (06-SEP-1999), Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, UK and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwenematode.wustl.edu
On Sep 5, 1999 this sequence version replaced gi:5763822.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
FEATURES
source
1. 254217
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="IV"

```

OM of: US-09-215-435-179 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:28 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_447 -NCPU=6 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179

Query length: 121

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	666.00	1331.22	3.5e-66	542
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	666.00	1328.81	4.8e-66	680
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	619.00	1238.97	4.9e-61	421
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	135.50	257.75	2.2e-06	656
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	135.50	257.59	2.2e-06	666
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	86.00	142.00	6.13	2898
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	86.00	138.62	9.47	3988
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	86.00	137.43	11.02	4460
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	84.50	154.32	1.26	683
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	137.27	11.25	2558
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	135.96	13.31	2895
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	135.87	13.46	2918
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	135.53	14.06	3014
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	135.43	14.25	3043
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	135.26	14.56	3092
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	135.11	14.84	3136
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	134.99	15.06	3170
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	134.72	15.60	3253
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	134.72	15.60	3253
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	134.71	15.62	3255
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	83.00	134.71	15.62	3255
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	83.00	134.71	15.62	3255
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	151.24	1.88	624
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	149.99	2.20	702
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	149.99	2.20	702
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	149.03	2.49	768
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	148.91	2.53	777
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	148.87	2.54	780
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	148.87	2.54	780
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	148.75	2.58	789
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	147.58	3.00	881
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	133.98	17.15	3170
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	133.98	17.15	3170
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.50	133.98	17.15	3170
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.50	132.35	21.13	3696
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.00	157.25	0.8675	322
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.00	137.95	10.31	1984
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97614	+	82.00	134.48	16.09	2751
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614	+	82.00	127.81	37.83	5154

/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA298069 + 81.50 146.83 3.30 781
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ37622 - 79.50 143.64 4.97 721
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA34914 + 79.00 145.60 3.86 545
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA208867 - 78.50 129.72 29.63 221
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AA794120 - 78.50 127.75 38.12 266

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97614

seq_documentation_block:

ID_AA97614 standard; DNA; 542 BP.

XX

AC_AA97614;

XX

DT 13-SEP-1999 (first entry)

XX

DE Extended human secreted protein coding sequence, SEQ ID NO. 78.

XX

KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
cellular differentiation; immune system regulator; anti-inflammatory;
haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
genetic disease; ss.

XX

OS Homo sapiens.

XX

PN WO9931236-A2.

XX

PD 24-JUN-1999.

XX

PF 17-DEC-1998; 98WO-IB02122.

XX

PR 10-AUG-1998; 98US-00961116.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0081563.

PA (GEST) GENSET.

PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
WPT; 1999-385906/32.

DR P-PSDB; AAY35930.

DR

PT New isolated human secreted proteins

XX

PS Claim 1; Page 212; 516pp; English.

XX

CC This sequence represents an extended human secreted protein coding
sequence of the invention. The secreted proteins can be used in treating
or controlling a variety of human conditions. The secreted proteins may
act as cytokines or may affect cellular proliferation or differentiation
or may act as immune system regulators, haematopoiesis regulators, tissue
growth regulators, regulators of reproductive hormones or cell movement
or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
tumour inhibition activity. The DNAs can be used in forensic procedures
to identify individuals or in diagnostic procedures to identify
individuals having genetic diseases resulting from abnormal expression of
the genes corresponding to the extended cDNAs. They are also useful for
constructing a high resolution map of the human chromosomes. They can
also be used for gene therapy to control or treat genetic diseases.

XX

SQ Sequence 542 BP; 132 A; 146 C; 126 G; 138 T; 0 other;

alignment_scores:
Quality: 666.00 Length: 121
Ratio: 5.504 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-179 x AA97614

Align seg 1/1 to: AA97614 from: 1 to: 542

1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuPhePheLeu 17
|||
16 ATCATGCTACCCCAATGGTGTCTGTCTTCCTTCTTTCTTTCT 65

17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34
|||||
66 CTTCTCTCTACCGGGGCTCACTTTCTCCAACAAAAATAAACCTTTGG 115
|||||
34 luLeuLysGluSerCyslleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
116 AGCTCAAGGAGTCTTGCATCCGGAACCAAGGACTCGAGACTGGCTGTC 165
|||||
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
166 CAACGTGCTCCACACANATGGCAGTCGCACCTCGCGGGAAGGGGTCGA 215
|||||
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
216 GGGCAGTCTGTGTCAACACAGGTGTTCTTTGGCCAATATAGACGGTGTC 265
|||||
84 roCysLeuArgAsnLeuThrCyslleTyrSerLysAsnGluLysTyrPLeu 100
|||||
266 CCTGCCCTGCCGACCTTGACTTGATATATATCAAAAGAAATGAGAAATGGCTT 315
|||||
101 SerlleAlaTyrGlyArgCysGlnLyslleGlyArgGlnLysLeuAlaLy 117
|||||
316 AGCATCGCCTATGGCGGTTGTCAAGAAATGGGAAGCGAAGTTGGCTAA 365
|||||
117 sLysMetPhePhe 121
|||||
366 CAAATATGTTCTTC 378

seq_documentation_block:	
AA26408 standard; cDNA; 680 BP.	
AAA26408;	
29-JUN-2000 (first entry)	
Human secreted protein gene 63 SEQ ID NO:73.	
Human; secreted protein; diagnosis: cytostatic; immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective; anti allergic; osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma; antispasmodic; cardiast; gene therapy; cancer; neurological disorder; immune disease; inflammation; blood disorder; tumour; ss.	
Homo sapiens.	
WO2000006598-A1.	
10-FEB-2000.	
29-JUL-1999; 99WO-US17130.	
30-JUL-1998; 98US-0094657.	
05-AUG-1998; 98US-0095486.	
06-AUG-1998; 98US-0095454.	
06-AUG-1998; 98US-0095455.	
12-AUG-1998; 98US-0096319.	
(HUMA-) HUMAN GENOME SCI INC.	
Komatsoulis GA, Rosen CA, Ruben SM, Moore PA, Shi Y;	
Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;	
Pi Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;	
WPI; 2000-195282/17.	
P-PSDB; AAY91513.	

117 sLysMetPhePhe 121
|||||
483 GAAATGTTCTTC 495

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAx51465

seq_documentation_block:

ID AAX51465 standard; cDNA; 421 BP.

XX AC AAX51465;

DT 21-JUN-1999 (first entry)

XX DE Human secreted protein 5' EST SEQ ID NO:44.

XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX OS Homo sapiens.

XX PN WO9906549-A2.

XX PD 11-FEB-1999.

XX PF 31-JUL-1998; 98WO-IB01231.

XX PR 01-AUG-1997; 97US-0905279.

XX PA (GEST) GENSET.

XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153779/13.

XX DR P-PSDB; AAY12687.

XX PT New nucleic acids encoding human secreted proteins - obtained from
XX CDNA libraries derived from testis, ovary, uterus and spleen tissue

XX PS Claim 1; Page 177; 522pp; English.

XX AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12681 to
XX AAY12913, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.

XX SQ Sequence 421 BP; 96 A; 106 C; 110 G; 102 T; 7 other;

alignment_scores:

Quality: 619.00 Length: 118
Ratio: 5.383 Gaps: 0
Percent Similarity: 97.458 Percent Identity: 95.763

alignment_block:

US-09-215-435-179 x AAX51465 ..

Align seg 1/1 to: AAX51465 from: 1 to: 421

1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuPhePhePhe 17
|||||
62 ATGATGCTACCCCAATGGCTGCTGCTGCTCTCTCTCTCTCT 111
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34
|||||
112 CTTCTCTCTCACCAGGGCTCACTTTCTCCACAAAAATACAACTTTTGG 161
34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
162 AGCTCAAGGAGSKKGCATSGGGAACACGAGACTGCGAGACTGGCTGCTGC 211
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
212 CACAGTGTCTCCACACAAATGCGAGTGCAGTGGCGGAGAGGGGTCGGA 261
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
262 GGGCACTCTGTGTCAAACGACGCTGTTCTTTGGCCAATATAGAGCGTGT 311
84 roCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
|||||
312 CCTGCCTGGGAACCTGACTTGTATATATTCAAAGAATGAGAAATGGCTT 361
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 117
|||||
362 AGCATCGCTATGCGCGTGTTCAGAAAAATGGAAAGGCAGAAATGGCTAG 411
117 sLys 118
:|||
412 RAAA 415

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ41401

seq_documentation_block:

ID AAZ41401 standard; cDNA; 656 BP.

XX AC AAZ41401;

XX DT 19-JAN-2000 (first entry)

XX DE Human normal pancreas tissue derived cDNA 16.

XX KW Human; pancreas; cancer; treatment; anticancer; cytostatic;
XX gene therapy; EST; expressed sequence tag; ss.

XX OS Homo sapiens.

XX PN DE19818598-A1.

XX PD 21-OCT-1999.

XX PF 19-APR-1998; 98DE-1018598.

XX PR 19-APR-1998; 98DE-1018598.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX DR WPI; 1999-592039/51.

XX PT New nucleic acid sequences expressed in normal pancreatic tissues, and
XX derived polypeptides, for treatment of pancreatic cancer and
XX identification of therapeutic agents

XX PS Claim 3; Page 56; 92pp; German.

XX CC This invention describes novel cDNA sequences (A) that are highly
XX expressed in normal human pancreatic tissue and which have anticancer

Thu Nov 15 10:52:00 2001

us-09-215-435-179.rng

CC and cytosolic activity. (A) are used (i) for recombinant expression of
CC polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to
CC identify agents suitable for treatment of pancreatic cancer: (ii)
CC directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AA241386-241423 represent human normal pancreatic
CC tissue derived cDNA fragments which encode the protein fragments
CC represented in AAV59893-Y59920.

XX Sequence 656 BP; 165 A; 200 C; 162 G; 129 T; 0 other;

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:

US-09-215-435-179 x AA241401 ..
Align seg 1/1 to: AA241401 from: 1 to: 656
12 LeuLeuPhePheLeuLeuLeuThrArgGlySerLeuSerProTh 28
98 ATCTGTATCCCTCCCTGCTGTCGCTCTCTGTCGCTATGCAGCTCTGG 147
28 rlystyrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
148 CCCCCGGGGATTTATCAACCTGGAGACGGTGTGCTGTGATGAATA 197
42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
198 GTGCCAGTGTAAAGACAAATGCTGCCAGCAATCAAGTGCCTGGGCTG 247
59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
248 GCCCGCTGCATCCATCCGCGCAGGAGAACGCGTGTCTGTCAAGAC 297
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
298 GCTCTATGGGATTTTACTACAGTGTCCCTGTGAGCGTGGCTGACCTGT. 346
92 leryrSerLysAsnGluLysTrpLeu.....SerIleAla 103
347GAGGGAGACAGACCATCGTGGGTCCATCACCACCAACAC 388
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
389 TTGGCATCTGCGCATGACGCTGCAGCTCCAAG 421

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC98928

seq_documentation_block:

ID AAC98928 standard; cDNA; 666 BP.
XX
AC AAC98928;
XX
DT 09-MAR-2001 (first entry)
XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:156.
DE
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
XX nototropic; diagnosis; identification; cytosolic; neuroprotective;
KW antinflammatory; immunomodulatory; relaxant; contraceptive; gynaecological;
KW linkage analysis; tissue identification; gene therapy; chromosome mapping;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW

pulmonary; cardiovascular; renal; proliferative; ss.

XX Homo sapiens.

XX WO200005320-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05989.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-579444/54.

XX P-PSDB; AAB54163.

XX New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 1; Page 614; 1379pp; English.

XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
XX AAB54466. The human pancreatic cancer antigens have cytosolic,
XX neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,
XX gynaecological, cardiant and antinflammatory activities, and can be used
XX in gene therapy. The polynucleotide and proteins can be used for
XX preventing, treating, or ameliorating a medical condition or in assays
XX for diagnosing a pathological condition or a susceptibility to one in a
XX subject. Binding partners to the proteins and the activity of the
XX proteins can be identified. The pancreatic cancer antigens can be used to
XX detect, treat or prevent pancreatic disorders, especially cancer.
XX Agonists and antagonists to the antigens can be screened for. The
XX pancreatic cancer antigen polynucleotides can be used to design nucleic
XX acid hybridisation probes that can be used in chromosome mapping, linkage
XX analysis, tissue identification and/or typing and a variety of forensic
XX and diagnostic methods. The proteins can be used to generate antibodies
XX which are used to purify, detect and target the polypeptides, including
XX both in vivo and in vitro diagnostic and therapeutic methods. The
XX proteins can be used to treat or prevent neural, immune system, muscular,
XX reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX sequences used in the exemplification of the present invention.

XX Sequence 666 BP; 151 A; 217 C; 157 G; 133 T; 8 other;

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:

US-09-215-435-179 x AAC98928 ..
Align seg 1/1 to: AAC98928 from: 1 to: 666
12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProTh 28
91 ATCTGTATCCCTCTGCTGTCGCTCTCTGTCGCTATGCAGCTCTCGG 140
28 rlystyrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
141 CCCCCGGGGATCATTTATCAACCTGGAGACGGTGTGCTGTGATGAATA 190
42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
191 GTGCCAGTGTAAAGACAAATGCTGCCAGCAATTCAGTGGCTGGGCTG 240

```

59 SerHisCysAlaGluLysGlySerGluGlySerLeuGlySerLeuGlnThrGlnVa 75
241 GCCCGCTGCACATCCATGCGCCAGCAGAGACGAGTGCTCTGTCAAGAC 290
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
291 GCTCTATGGGATTACTACAAAGTGCCCTGTGAGCGTGCCCTGAGACTGT. 339
92 letyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
340 .....GAGGAGACACAGACCATCGTGGGCTCCATCACCACCAAC 381
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
382 TTTGGCATCTGCATGACGCTGGAGGCTCCAAG 414

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT.AAV35367

seq_documentation_block:

ID AAV35367 standard; cDNA; 2898 BP.

AC AAV35367;

DT 01-OCT-1998 (first entry)

DE Human semaphorin encoding cDNA.

KW Human; semaphorin; diagnosis; nerve disease; nerve growth inhibitor;
neurological disease; atopic skin inflammation; autoimmune disease;
pain; ds.

OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 370..2897

FT /*tag= a

/product= "semaphorin"

PN W09822504-A1.

PD 28-MAY-1998.

PF 12-NOV-1997; 97WO-JP04111.

PR 15-NOV-1996; 96JP-0321068.

PA (SUMU) SUNITOMO PHARM CO LTD.

PI Furuyama T, Inagaki S;

DR WPI: 1998-312416/27.

DR P-PSDB; AAW63748.

PT Gene encoding new semaphorin nerve growth inhibitor - useful in
diagnosis, treatment and study of neurological diseases

PS Claim 2; Page 37-39; 49pp; Japanese.

CC The present sequence encodes human semaphorin, a nerve growth
inhibitor. The semaphorin protein, and gene encoding the protein,
and their derivatives, are used in the diagnosis, treatment and
study of neurological disorders such as atopic skin inflammation,
autoimmune diseases and pain.

SQ Sequence 2898 BP; 794 A; 695 C; 748 G; 660 T; 1 other;

alignment_scores:

Quality:	86.00	Length:	87
Ratio:	1.536	Gaps:	7
Percent Similarity:	64.368	Percent Identity:	35.632

alignment_block:

```

US-09-215-435-179 x AAV35367/rev ..
Align seg 1/1 to reverse of: AAV35367 from: 1 to: 2898

16 PheLeuPheLeuLeuThrArgGlySerLeuSerProThrLysTyr.AsnL 32
1103 TCATTTTGTTCATCTCGGTCTTTCAGGAGCGCTCATCGTCATATATGAACC 1054
32 eu.....LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGlu 46
1053 TACAATTTTGGTTCTTTCAGGAGCGCTCATCGTCATCTCAGTCCGAA 1004
47 ThrGly.....CysCysGlnArgAlaProAspAsnCys...G1 58
1003 TATGGCTAACTTCCCATGCTGCGGAAGATCGCGAGTCTCTGCCCAA 954
58 userHisCys...AlaGluLysGlySerGluGlySerLeuGlySerGlnThrG 74
953 TAGTCACTAGTACGTCACGCAACACAGCTCATTTCCCAACTAGCGTGGACAC 904
74 In...ValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeu 89
903 AAGGAGGAGTGGGGTCAAAAGGACATCTGCCCCCTCTCTCTCAGATC 854
90 ThrCysIle 92
853 ...TGTGTG 848

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT.AAZ28470

seq_documentation_block:

ID AAZ28470 standard; DNA; 3988 BP.

AC AAZ28470;

DT 05-JAN-2000 (first entry)

DE Mouse semaphorin H variant (Sema Hv) polynucleotide sequence.

KW Semaphorin H; Sema H; Sema Hv; collapsin; bone structure formation;
metastasis; cancer; antibody; drug screen; ss.

OS Mus sp.

FH Key Location/Qualifiers
FT CDS 609..2943

FT /*tag= a

FT /product= Sema_Hv

/note= "Mouse semaphorin H variant"

PN W09947671-A2.

PD 23-SEP-1999.

PF 12-MAR-1999; 99WO-IB00495.

PR 13-MAR-1998; 98US-0077997.

PA (LUKA/) LUKANIDIN E M.

PA (CHRI/) CHRISTENSEN C R L.

PI Lukanidin EM, Christensen CRL;

DR WPI: 1999-590975/50.

DR P-PSDB; AAY43091.

PT New polypeptides and polynucleotides, useful in diagnosis and treatment
of metastatic cancer

PS Claim 1; Page 85-91; 95pp; English.

CC This is the mouse Semaphorin-H variant (Sema Hv) polynucleotide sequence
(Sema Hv). The Semaphorin/Collapsin family of molecules are characterised

XX	Mus sp.	
OS		
XX	Location/Qualifiers	
FFH	609..2937	
CDS	/*tag= a	
FT	/product= M_Sema-H	
FT	/trans_except=(Pos:2452..2454, aa:Tyr)	
FT		
XX		
PN	W09947671-A2.	
XX		
PD	23-SEP-1999.	
XX		
PF	12-MAR-1999; 99WO-IB00495.	
XX		
PR	13-MAR-1998; 98US-0077997.	
XX		
PA	{LUKA/) LUKANIDIN E.M.	
PA	(CHRI/) CHRISTENSEN C R L.	
PI	Lukanidin EM, Christensen CRL;	
XX		
XX	WPI; 1999-590975/50.	
DR	P-PSDB; AA228469.	
DR		
XX	New polypeptides and polynucleotides, useful in diagnosis and treatment	
PT	of metastatic cancer	
PT		
XX	Claim 1; Page 78-85; 95pp; English.	
XX		
CC	This is the mouse Semaphorin-H polynucleotide sequence (Sema H). The	
CC	Semaphorin/collapsin family of molecules are characterised by a unique	
CC	and highly conserved motif, within a 500 amino acid semaphorin domain.	
CC	Some semaphorins exhibit inhibitory or repulsive functions in a neuronal	
CC	context, and functions in bone structure formation are also implicated.	
CC	It is thought that Sema H plays a role in metastasis. The invention uses	
CC	the mouse Sema H gene in the diagnosis of metastatic cancer. Semaphorin	
CC	polynucleotides and polypeptides are used in the methods of the	
CC	invention, the polypeptides are useful for determining the metastatic	
CC	potential of cells, by detecting their expression in biological samples.	
CC	Antibodies specific for Sema H, are also useful therapeutically in	
CC	inhibiting Sema-H polypeptide activity and therefore metastasis, and for	
CC	purifying the polypeptides. Metastasis may also be inhibited by	
CC	inhibiting the biological activity of the polypeptide using e.g. a small	
CC	molecule inhibitor or a Semaphorin-H ligand (or fragment). The	
CC	polynucleotides can also be used to inhibit polypeptide expression in	
CC	cells using known antisense technology e.g. to prevent metastasis of	
CC	cancer cells. They can be used to detect and quantify Sema-H mRNA levels	
CC	in cells. The polypeptides, fusion proteins, multimeric proteins,	
CC	antibodies or antisense oligonucleotides can be included in	
CC	pharmaceutical compositions. The polynucleotides can be used to isolate	
CC	similar sequences from other species and to produce mammalian cell lines	
CC	and tumours with known metastatic potential, useful in anti-metastatic	
CC	drug screening.	

```

alignment_scores:      86.00      Length:      87
                      Ratio: 1.536      Gaps:      7
                      Percent Similarity: 64.368      Percent Identity: 35.632

```

alignment_block: 00-215-425-170 v AAZ28469/rev

US-09-215-433-173 & AM2001027, E.
 173 to number of: A9728469 from: 1 to: 4460

16 PheLeuPheLeuLeuThrArgGlySerLeuSerProThrLysTyr AsnL 32
::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
1343 TACATTTTGTCATCTCGGTTCGGTTATCAGGAATCATATCAACC 129

32 en LeuGlutLeuLysGluSerCysIleArgAsnGlnAspCysGlu 46


```

CC      method of the invention.
XX
SQ      Sequence 683 BP; 198 A; 133 C; 127 G; 225 T; 0 other.

alignment_scores:
    Quality:      84.50      Length:      99
    Ratio:        1.760     Gaps:        4
    Percent Similarity: 48.485      Percent Identity: 30.303

alignment_block:
    US-09-215-435-179 x AAF74198      ..
Align seq 1/1 to: AAF74198 from: 1 to: 683

```

Align seq 1/1 to: AAF74198 from: 1 to: 683

116 AGTTGAACCTATTTTCATCCCT

TTG AGT TGA ACC TAT T T C AT C C C A A T C A A G T G G G T A G C A F C A C T G G A A G T A 103

cc
eulys

100 CCICAGAGTAAAGTG.....GATTTGGGGCCGACGTACGGTG 200

02 ARGATA.....FTOASPRASINUSGRUSETNISCYATAGT 03

||||: |||:: |||:: |||:: |||::

Z07 AUG96CCICGAAAGGGCCAAAACCAATGCGAACAAGGTCATGT...Z08

[illegible][illegible][illegible][illegible]

```
seq_name: /cgr2-/gcgaaca/geneseq/na2001.dat:chr2243
```

seq_documentation_block.
ID AAF55249 standard; cDNA; 2558 BP.

AC AAF55249;

DT 29-MAY-2001 (first entry)

DE Nucleotide sequence of murine parkin2 with a frameshift mutation

KW Parkin2; parkinson's disease; 6q25.2-27; neurodegenerative

KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis; Wilson's disease; Pick's disease; Prion protein; Multi-system atrophy; Wilson's disease; Pick's disease; Prion protein

KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity; ss.

OS Mus sp.

aa	FH	Key	Location/Qualifiers

```

E1      CUS
E2      129.1320
E3      /*taq= a

```

XX
17
-XXXXXXXXXXXX-
XXXXXXXXXXXX

[illegible]

FD
XX
07 MAR 2001.

XX
00 AUG 1955, 0010700.

30 AUG 1993, 22EF 0110700.
 FK XX

XX
XX
(PIOT) PIOTKONILENS FIANT OMBH.

[illegible]


```
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55253
seq_documentation_block:
ID  AAF55253 standard: cDNA; 3043 BP.
XX
XX  AAF55253;
AC  AAF55253;
XX
XX  29-MAY-2001 (first entry)
DT
XX
XX  Nucleotide sequence of murine parkin2 with a frameshift mutation.
DE
XX
XX  Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW  Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW  Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW  brain tumour; head trauma; stroke; vascular irregularity;
KW  metabolic irregularity; ss.
XX
XX  Mus sp.
OS
XX
XX  Key Location/Qualifiers
FH  129..1019
FT  CDS
FT  /*tag= a
FT  /product= "truncated parkin2"
XX
XX  EP1081225-AL.
PN
XX  07-MAR-2001.
PD
XX  30-AUG-1999; 99EP-0116766.
PF
XX  30-AUG-1999; 99EP-0116766.
PR
XX  (BIOF-) BIOFRONTERA PHARM GMBH.
PA
XX  Luebbert H;
PI
XX
XX  WPI; 2001-212797/22.
DR  P-PSDB; AAB67526.
XX
XX  New polynucleotides encoding mouse parkin2 protein, useful for
XX  producing a transgenic non-human animal as an animal model for
XX  neurodegenerative diseases
XX
XX  Claim 3; Page 29-30; 62pp; English.
XX
XX  The present sequence encodes a murine parkin2 polypeptide. The
XX  polynucleotide sequence contains a frameshift mutation, leading to a
XX  truncated protein. Mutations or deletions in the parkin2 gene cause
XX  Parkinson's disease in humans. The human parkin2 gene is located in
XX  gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are
XX  useful for analysing neurodegenerative diseases. They are also useful
XX  for testing the efficacy of the treatment of a neurodegenerative disease
XX  such as Parkinson's disease, Alzheimer's disease, Huntington's disease,
XX  amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
XX  Pick's disease, Prion disease, and secondary causes inducing
XX  Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
XX  stroke, vascular irregularities or metabolic irregularities, associated
XX  with a less active or non-active parkin protein.
XX
XX  Sequence 3014 BP; 781 A; 761 C; 734 G; 738 T; 0 other;

alignment_scores:
Quality: 83.00 Length: 106
Ratio: 1.566 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 29.245

alignment_block:
US-09-215-435-179 x AAF55247 ..
Align seg 1/1 to: AAF55247 from: 1 to: 3014
6 TrpLeuLeuLeuLeuPheLeuLeuPhePheLeuLeuLeuThrAr 22
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2332 TTTTATTGTTTGGCTTTGCTTTTGGCTTTTGGCTTTTGGCTTTG 2381
22 gGly.....SerLeuSerProThrLysTyrAsnLeuLeuLeuLysG 37
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2382 GGGGCTACCAACTTGAGGCCTTAGAGCTAAATCATATAGATGATGT 2431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
37 lUsrCysIleArgAsnGlnAspCysGluThrGlyCysGlnArgAla 53
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2432 TATCTGTGG.....TGTGAGGAAGG 2453
54 ProAspAsnCysGluSerHisCysAlaGluLysGly..... 65
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2454 CCAGCTGGCCTAAGTCACACTTTTGTCCAGTGCCCTAGACTCCACCC 2503
66 .....SergluGlySerL 70
2504 AGCCAGCTCCCAAAATGAAAGACCACTGTCAAGCAGCAGTCAGGAGTC 2553
70 eu.CysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysProCysLe 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2554 TGATGTCACCCATCACTATTTTTCATCATGTGCTTGCCTCTGCT 2603
86 uArgAsnLeuThrCys 91
|||
2604 CCTTCCACACCCGCTGT 2619
```


us-09-215-435-179.rng

Thu Nov 15 10:52:00 2001

Thu Nov 15 10:52:00 2001

```

; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,740
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2663 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 139..2049
; US-08-653-740-4

alignment_scores:
  Quality: 78.50 Length: 125
  Ratio: 1.287 Gaps: 8
  Percent Similarity: 48.800 Percent Identity: 28.800

alignment_block:
  US-09-215-435-179 x US-08-653-740-4 rev ..
  Align seg 1/1 to reverse of: US-08-653-740-4 from: 1 to: 2663

4 ProGlnTrpLeu.....LeuLe 9
|||||
2533 CCCAAGTGGCTCGGACTACAGGTGTGTGCCATCATGCCAGCTAATTGTT 2484

9 uLeuPheLeuLeuPhePheLeuPheLeuLeuThrArgGlySer.... 24
|||||
2483 TTGTGTTTGTGTTTATTT.....TGTAGTGAGGGGTCTTGCT 2443

25 ..LeuSerProThrLysTyrAsnLeuLeuLeuLeuLysGluSerCysIle 40
|||||
2442 ATATTGCCAGGCTGGTCTCAAACTCTAGCTCTA..... 2408

41 ArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCy 57
|||||
2407 .....AGTGATCCTTCCACCTGGCTGGCTGCCAAGTGTGGATTACAGG 2364

57 sGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrG 74
|||||
2363 CGTGAGCCACTGTGCT.....GGGGTCCAAATTCAGTCTCTGTACAGC 2320

74 InValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThr 90
|||||
2319 GAACAGATTGTC.....AGGAGCTGCCACTC.....TCAATTCA 2282

```

```

91 CysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgCy 107
|||||
2281 TGTATTTTC.....CTTGTCTCCACCTTCATCCGCTG 2250

107 sGlnLysIleGlyArgGlnLysLeu 115
|||||
2249 TCYTCCTCCATGGGGCCCAAGCCCTT 2225

seq_name: /cgnl_7/ptodata/1/lna/5B_COMB.seq:US-09-073-594-4

seq_documentation_block:
; Sequence 4, Application US/09073594
; Patent No. 5925735
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,594
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2663 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 139..2049
; US-09-073-594-4

alignment_scores:
  Quality: 78.50 Length: 125
  Ratio: 1.287 Gaps: 8
  Percent Similarity: 48.800 Percent Identity: 28.800

alignment_block:
  US-09-215-435-179 x US-09-073-594-4 rev ..
  Align seg 1/1 to reverse of: US-09-073-594-4 from: 1 to: 2663

4 ProGlnTrpLeu.....LeuLe 9
|||||
2533 CCCAAGTGGCTCGGACTACAGGTGTGTGCCATCATGCCAGCTAATTGTT 2484

9 uLeuPheLeuLeuPhePhePheLeuPheLeuLeuThrArgGlySer.... 24
|||||
2483 TTGTGTTTGTGTTTATTT.....TGTAGTGAGGGGTCTTGCT 2443

```



```
25 ..LeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysIle 40
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2442 ATATTGCCAGCGCTGCTCAAACTCCTAGCTCTA..... 2408

41 ArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCy 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2407 .....AGTGATCCTCCACCTGGCTGCCAAAGTGCTGGATTACAGG 2364

57 sGluSerHisCysAlaGluLysGlySerGluGlySerLeuGluThrG 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2363 CGTGAGCCACCTGTGCT.....GGGTCCAAATTCAGTCTCTTACAGG 2320

74 InValPheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThr 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2319 GAACAGATTTGGC.....AGGCAGCTGCCACTC.....TCAATTCA 2282

91 CysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgCy 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2281 TGTATTTTC.....CTTTGCTCCACCTCCATCCGCG 2250

107 sGlnLysIleGlyArgGlnLysLeu 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2249 TCTTCCCCATGGGGGCCAAGCCCTT 2225
```

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-275-925-4

```
seq_documentation_block:
; Sequence 4, Application US/09275925
; Patent No. 6080406
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/275,925
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2663 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 139..2049
US-09-275-925-4
```

```
alignment_scores:
Quality: 78.50 Length: 125
Ratio: 1.287 Gaps: 8
Percent Similarity: 48.800 Percent Identity: 28.800

alignment_block:
US-09-215-435-179 x US-09-275-925-4/rev ..
Align seg 1/1 to reverse of: US-09-275-925-4 from: 1 to: 2663

4 ProGlnTrpLeu.....LeuLe 9
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2533 CCCAAGTGGCTCGGACTACAGGTGTGTGCCATCATGCCAGCTAATTGTT 2484

9 uLeuPheLeuPhePhePheLeuPheLeuLeuThrArgGlySer.... 24
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2483 TTTGTTTTTGAITTTTATTTT.....TGTAGTGAGGGGTCTTGCT 2443

25 ..LeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysIle 40
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2442 ATATTGCCAGCGCTGCTCAAACTCCTAGCTCTA..... 2408

41 ArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCy 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2407 .....AGTGATCCTCCACCTGGCTGCCAAAGTGCTGGATTACAGG 2364

57 sGluSerHisCysAlaGluLysGlySerGluGlySerLeuGluThrG 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2363 CGTGAGCCACCTGTGCT.....GGGTCCAAATTCAGTCTCTTACAGG 2320

74 InValPheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThr 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2319 GAACAGATTTGGC.....AGGCAGCTGCCACTC.....TCAATTCA 2282

91 CysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgCy 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2281 TGTATTTTC.....CTTTGCTCCACCTCCATCCGCG 2250

107 sGlnLysIleGlyArgGlnLysLeu 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2249 TCTTCCCCATGGGGGCCAAGCCCTT 2225

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-085-199B-3

seq_documentation_block:
; Sequence 3, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
; APPLICANT: Hayden. Michael R.
; APPLICANT: Hackam, Abigail
; APPLICANT: Huq, A.H.M. Mahbubul
; APPLICANT: Chopra, Vikramjit Singh
; APPLICANT: Kalchman, Michael
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
; TITLE OF INVENTION: Huntington's Disease Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppedahl & Larson
; STREET: PO Box 5270
; CITY: Frisco
; STATE: CO
; COUNTRY: USA
; ZIP: 80443-5270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS DOS 5.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,199B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```



```
82 .....AlaCysProCysLeu 86
60 ATGAGCTAGCGTGCCCGTGCCTA 38

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-357-141-63

seq_documentation_block:
; Sequence 63, Application US/09357141
; Patent No. 6277825
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: McCabe, R. Tyler
; APPLICANT: Layer, Richard T.
; APPLICANT: Zhou, Li-Ming
; TITLE OF INVENTION: Use of Conantokins for Treating Pain
; FILE REFERENCE: 2314-171
; CURRENT APPLICATION NUMBER: US/09/357,141
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 09/283,277
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: US 09/142,078
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: WO US97/12652
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: US 08/762,377
; PRIOR FILING DATE: 1996-12-06
; PRIOR APPLICATION NUMBER: US 08/684,750
; PRIOR FILING DATE: 1996-07-22
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Conus characteristic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(279)
US-09-357-141-63

alignment_scores:
Quality: 77.50 Length: 91
Ratio: 1.550 Gaps: 4
Percent Similarity: 54.945 Percent Identity: 29.670

alignment_block:
US-09-215-435-179 x US-09-357-141-63/rev ..

Align seg 1/1 to reverse of: US-09-357-141-63 from: 1 to: 356

11 PheLeuLeuPhePhePheLeuLeuThrArgGlySer...LeuSe 26
||| :|||:|||||:|||||: ||| ||
298 TTTAGAACTATCTTTTCTATTCTCTCGGTTCTCGAGTCTCTTAACAGTC 249
||| :|||:|||||:|||||: ||| ||
26 rProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysIle..... 40
||| :|||:|||||:|||||: ||| ||
248 TCGGCAATCTCTATCTTCTCTCGTAGCCCTCGGCATGTTTCTCGCTT 199
||| :|||:|||||:|||||: ||| ||
41 .....ArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAsp 55
:|||||:|||||: |||: ||| |||
198 TTTGAGAGTCTCTCTCTGAGTCAACCTGT.....CCTTGC 161
:|||||:|||||: |||: ||| |||
56 AsnCysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysG1 72
:|||||:|||||: |||: ||| |||
160 CATTTGCTGTCAGTGTGCGGGCAGCGGATTTCTGCAGGAGGACAGGCTCA 111
:|||||:|||||: |||: ||| |||
72 nThrGlnValPhePheGlyGlnTyrArg..... 81
:|||||:|||||: |||: ||| |||
110 GGTTCAGTGTGTCGATCATGCCGACCGCGGTTCAGTCAGTGGCTCTCC 61
:|||||:|||||: |||: ||| |||
82 .....AlaCysProCysLeu 86
```

```
60 ATGAGCTAGCGTGCCCGTGCCTA 38

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-11

seq_documentation_block:
; Sequence 11, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136..2543
; US-07-862-021B-11

alignment_scores:
Quality: 76.00 Length: 101
Ratio: 1.407 Gaps: 4
Percent Similarity: 53.465 Percent Identity: 21.782

alignment_block:
US-09-215-435-179 x US-07-862-021B-11 ..

Align seg 1/1 to: US-07-862-021B-11 from: 1 to: 3226

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
||| :|||:|||||:|||||: ||| |||
1747 GAAGAACCTGAGAAATGTATTGTAATGAGGAATGCTCCCTAGCAGCTG 1796
:|||||:|||||: |||: ||| |||
50 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAlaG 63
:|||||:|||||: |||: ||| |||
1797 CCTTGTCCACCGAATGGGAGAGTGGGATGAATGCAGTGTAGCTGTGGCA 1846
:|||||:|||||: |||: ||| |||
63 lu.....LysGlySerGluGly 68
:|||||:|||||: |||: ||| |||
1847 CAGGAATGAAAGGCGACACAGAATGATCAAGATGACTCTCTGTATGGA 1896
:|||||:|||||: |||: ||| |||
```

seq name: /cqn1.7/ptodata/1/ina/5A_COMB.seq:US-08-313-288B-11

Align seq 1/1 to: US-08-313-288B-11 from: 1 to: 3226

Thu Nov 15 10:52:00 2001

us-09-215-435-179.rni

```
;
; REFERENCE/DOCKET NUMBER: 40399/304/NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1556 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; STRAIN: Human Dlk
;
US-08-457-135-3

alignment_scores:
  Quality: 75.00      Length: 61
  Ratio: 2.143       Gaps: 5
  Percent Similarity: 57.377      Percent Identity: 34.426

alignment_block:
US-09-215-435-179 x US-08-457-135-3 ..

Align seg 1/1 to: US-08-457-135-3 from: 1 to: 1556

35 LeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysGlu 51
   ||| : : : : : ||| : : : : : |||
330 CTTGTGACAGTGGCGTGACCTCTCCGGCTGCTTCACGAGCTCTGTGG 379

51 nArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerGlu 68
   : : : : : ||| : : : : : ||| : : : : :
380 AGAA...CCGGGCGAGTGC.....ATTTCACCCGAC...GGCTGGGACG 417

68 lysrLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys... 83
   ||| : : : : : ||| : : : : : ||| : : : : :
418 GGGAGCTCTGTGATAGAGATGTT.....CGGGCTGCTCTCC 452

84 .....ProCysLeuArgAsnLeuThrCysIle 92
   ||| : : : : : ||| : : : : : ||| : : : : :
453 TCGGCCCCCTGTGCCAACACAGCGGACCTGGGTG 485

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-882-046-3

seq_documentation_block:
; Sequence 3, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
;
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16...3460
;
US-08-882-046-3

alignment_scores:
  Quality: 75.00      Length: 99
  Ratio: 1.531       Gaps: 4
  Percent Similarity: 49.495      Percent Identity: 23.232

alignment_block:
US-09-215-435-179 x US-08-882-046-3 ..

Align seg 1/1 to: US-08-882-046-3 from: 1 to: 4315

31 AsnLeuLeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluTh 47
   ||| : : : : : ||| : : : : : |||
892 AACGTGTGAAGGCTGAGCAGCGCTGCACCTCCAAACCCGTGTGCCAACGG 941

47 rGlyCysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGlu 64
   ||| : : : : : ||| : : : : : ||| : : : : :
942 GGGCTCTTGCATGAGTGGCTCGCGCTTCGAATGCCACTGC...CCAT 988

64 ysglySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyr 80
   ||| : : : : : ||| : : : : : ||| : : : : :
989 CGGGCTGGAGCGGCCACCTGTGCCCTTGACATC.....GATGAGTGT 1032

81 ArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsn... 96
   : : : : : ||| : : : : : ||| : : : : :
1033 GTTTCGAACCCGTGTGCGCGCGTGGCACCTGTGTGGACACGAGTGGACGG 1082

97 .....GluLysTrpLeuSerIle..... 102
   ||| : : : : : ||| : : : : : ||| : : : : :
1083 CTTTGAGTGCATCTGCCCGGACGAGTGGTGGGCCACCTGCCACGCTGG 1132

103 .....AlaTyrGlyArgCysGlnLysIleGlyArgGlnLys 114
   : : : : : ||| : : : : : ||| : : : : :
1133 ACGTCAACGACTGTGCGGGCAGTGTACGATGGGGGGCACCCTGCAAG 1179

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-181-706-1

seq_documentation_block:
; Sequence 1, Application US/09181706
; Patent No. 6130068
; GENERAL INFORMATION:
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,
; APPLICANT: Robert F. Dubose, Richard S. Johnson
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/181,706
; FILING DATE: October 28, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/958,598 (converted to a
; APPLICATION NUMBER: Provisional, see below)
; FILING DATE: October 28, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598
; APPLICATION NUMBER: conversion to Provisional application)
; FILING DATE: October 26, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; US-09-181-706-1

alignment_scores:
  Quality: 75.00 Length: 78
  Ratio: 1.705 Gaps: 7
Percent Similarity: 56.410 Percent Identity: 33.333

alignment_block:
US-09-215-435-179 x US-09-181-706-1 ..
Align seg 1/1 to: US-09-181-706-1 from: 1 to: 4707

56 AsnCysGlu.....SerHisCysAlaGluLysGlySerGl 67
:::||||| |:::|:::|:::|:::|
1296 TCCTGTGAAGAATATCTACATTTATCTAAACAGCTGGGAAGAGGTGAGCA 1345

67 uGlySerLeuCysGlnThrGln.....ValPhePheGlyGlnTyr. 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1346 GAATTCGTGTGCACAACTGCAATAACATAAATCCTGTCGGAGGTGTTTA 1395

81 .....ArgAla.CysProCysLeuArgAsnLeuTh 90
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1396 ACAGCCACAGACCTCAGTCGGTGGTGGCCATTGCTACAAAGAGTGCAC 1445

90 r.....CysIleTyrSerLysAsn...GluLysTrpLeuSerI 102
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1446 TTTTCAAGGAGATTGTGTACATTTCAGAGAACITAGAAAACTGGCTGGATA 1495

102 leaLatyrGly.....ArgCysGlnLysIle 110
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1496 TTTTCGTCTGGGCAAAAAGTGCCTAAAATT 1527

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-458-791-1

seq_documentation_block:
; Sequence 1, Application US/09458791
; Patent No. 6174689
; GENERAL INFORMATION:

```

```

; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/458,791
; FILING DATE: 10-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-458-791-1

alignment_scores:
  Quality: 75.00 Length: 78
  Ratio: 1.705 Gaps: 7
Percent Similarity: 56.410 Percent Identity: 33.333

alignment_block:
US-09-215-435-179 x US-09-458-791-1 ..
Align seg 1/1 to: US-09-458-791-1 from: 1 to: 4707

56 AsnCysGlu.....SerHisCysAlaGluLysGlySerGl 67
:::||||| |:::|:::|:::|:::|
1296 TCCTGTGAAGAATATCTACATTTATCTAAACAGCTGGGAAGAGGTGAGCA 1345

67 uGlySerLeuCysGlnThrGln.....ValPhePheGlyGlnTyr. 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1346 GAATTCGTGTGCACAACTGCAATAACATAAATCCTGTCGGAGGTGTTTA 1395

81 .....ArgAla.CysProCysLeuArgAsnLeuTh 90
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1396 ACAGCCACAGACCTCAGTCGGTGGTGGCCATTGCTACAAAGAGTGCAC 1445

90 r.....CysIleTyrSerLysAsn...GluLysTrpLeuSerI 102
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1446 TTTTCAAGGAGATTGTGTACATTTCAGAGAACITAGAAAACTGGCTGGATA 1495

102 leaLatyrGly.....ArgCysGlnLysIle 110

```

us-09-215-435-179.rni

Thu Nov 15 10:52:00 2001

|||||
1496 TTTCGTCTGGAGCAAAAAGTGCCTAAATT 1527

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgnl_7/ptdata/1/pna/US0902_COMB.seq	US-09-215-435-78	+666.00	1253.41	1.8e-60	542
/cgnl_7/ptdata/1/pna/US0909_COMB.seq	US-60-096-116-61	+666.00	1252.98	1.9e-60	565
/cgnl_7/ptdata/1/pna/PCTUS_COMB.seq	PCT-US99-17130-73	+666.00	1251.08	2.4e-60	680
/cgnl_7/ptdata/1/pna/US0904_COMB.seq	US-09-489-847-73	+666.00	1251.08	2.4e-60	680
/cgnl_7/ptdata/1/pna/US0918_COMB.seq	US-09-834-366-4817	+661.00	1244.08	5.9e-60	533
/cgnl_7/ptdata/1/pna/US0019_COMB.seq	US-60-197-873-4817	+661.00	1244.08	5.9e-60	533
/cgnl_7/ptdata/1/pna/US0899_COMB.seq	US-08-905-279-44	+619.00	1166.77	1.2e-55	421
/cgnl_7/ptdata/1/pna/US0881_COMB.seq	US-08-196-4814-520	+602.00	1132.98	9.2e-54	488
/cgnl_7/ptdata/1/pna/US0881_COMB.seq	US-08-196-4814-520	+602.00	1132.98	9.2e-54	488
/cgnl_7/ptdata/1/pna/US0881_COMB.seq	US-08-196-4814-520	+602.00	1132.98	9.2e-54	488
/cgnl_7/ptdata/1/pna/US0801_COMB.seq	US-60-184-698-48	+497.00	934.19	1.1e-42	463
/cgnl_7/ptdata/1/pna/US0618_COMB.seq	US-60-184-698-694	+497.00	933.86	1.1e-42	478
/cgnl_7/ptdata/1/pna/PCTUS_COMB.seq	PCT-US99-4430-3137-1	+497.00	932.94	1.3e-42	523
/cgnl_7/ptdata/1/pna/US0904_COMB.seq	US-09-489-847-73	+497.00	932.94	1.3e-42	523
/cgnl_7/ptdata/1/pna/US0907A_COMB.seq	US-09-713-067-1	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0907B_COMB.seq	US-09-731-872-89	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0908_COMB.seq	US-09-834-366-5913	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0908_COMB.seq	US-09-876-997-89	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0014_COMB.seq	US-60-147-499-1374	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0616_COMB.seq	US-60-169-629-108	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0618_COMB.seq	US-60-187-470-108	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US06019_COMB.seq	US-60-197-873-5913	+383.50	717.65	1.2e-30	514
/cgnl_7/ptdata/1/pna/US0909_COMB.seq	US-08-992-332-1466	+383.00	723.58	5.8e-31	263
/cgnl_7/ptdata/1/pna/US06003_COMB.seq	US-60-033-364-14	+383.00	723.58	5.8e-31	263
/cgnl_7/ptdata/1/pna/US06021_COMB.seq	US-60-212-664-192	+346.00	586.81	2.4e-23	171
/cgnl_7/ptdata/1/pna/US06021_COMB.seq	US-60-212-664-192	+346.00	586.81	2.4e-23	171
/cgnl_7/ptdata/1/pna/US06021_COMB.seq	US-60-212-664-192	+346.00	586.81	2.4e-23	171
/cgnl_7/ptdata/1/pna/US0959A_COMB.seq	US-09-528-409-86528	+188.00	347.96	4.9e-10	44
/cgnl_7/ptdata/1/pna/US0909_COMB.seq	US-09-933-524-86528	+188.00	347.96	4.9e-10	44
/cgnl_7/ptdata/1/pna/US0902_COMB.seq	US-09-274-861-9011	+135.50	249.49	0.0001	397
/cgnl_7/ptdata/1/pna/US0909_COMB.seq	US-09-515-738-9011	+135.50	249.49	0.0001	397
/cgnl_7/ptdata/1/pna/US0922_COMB.seq	US-09-284-697A-15	+135.50	246.66	0.0002	523
/cgnl_7/ptdata/1/pna/US096C_COMB.seq	US-09-654-617-289898	+135.50	245.69	0.0002	523
/cgnl_7/ptdata/1/pna/US096E_COMB.seq	US-09-684-016-289898	+135.50	245.69	0.0002	523
/cgnl_7/ptdata/1/pna/US06021_COMB.seq	US-60-213-360-121	+135.50	245.43	0.0003	590
/cgnl_7/ptdata/1/pna/US06032_COMB.seq	US-60-324-185-654	+135.50	245.43	0.0003	590
/cgnl_7/ptdata/1/pna/US096C_COMB.seq	US-09-652-914-7251	+135.50	245.25	0.0003	666
/cgnl_7/ptdata/1/pna/US096C_COMB.seq	US-09-647-801-16	+135.50	244.34	0.0003	656
/cgnl_7/ptdata/1/pna/PCTUS_COMB.seq	PCT-US00-05989-156	+135.50	244.18	0.0003	666
/cgnl_7/ptdata/1/pna/US0909_COMB.seq	US-09-925-297-156	+135.50	244.18	0.0003	666
/cgnl_7/ptdata/1/pna/US096C_COMB.seq	US-09-652-914-8817	+135.50	243.03	0.0003	740

```

166 CAACGTGCTCCACACANTTGCAGTCGCAGTCGCCGGAGAGGGGTCCGA 215
167 uGlySerLeuCySGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
216 GGGCAGCTCTGTGTCAAACGACGAGGTGTCTTTGGCCAAATATAGAGCGTGC 265
84 rOCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTirLeu 100
266 CCGTCTGGCGGAACCTTGACTTGATATATTTCAAGCAANTGAGNAATGGCTT 315
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaLy 117
316 AGCATCGCTATGGCCGTGTTCAGAAATTTGGAAGGCAGAGTTGGCTAA 365
117 sLysMetPhePhe 121
366 GAAATGTTCTTC 378

seq_name: /cgnl_7/ptodata/1/pna/US6009_COMB.seq:US-60-096-116-61
seq_documentation_block:
; Sequence 61, Application US/60096116.
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SECRETED PROTEINS
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/096,116
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.038PR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 16..84
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 9.8
; OTHER INFORMATION: seq FLFFFLLFLTRG/SL
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 502..507
; FEATURE:
; NAME/KEY: poly_a
; LOCATION: 529..542

```

102 EARLIER FILING DATE: 1998-08-05
103 EARLIER APPLICATION NUMBER: 60/095,455
104 EARLIER FILING DATE: 1998-08-06
105 EARLIER APPLICATION NUMBER: 60/095,454
106 EARLIER FILING DATE: 1998-08-06
107 EARLIER APPLICATION NUMBER: 60/096,319
108 EARLIER FILING DATE: 1998-08-12
109 NUMBER OF SEQ ID NOS: 364
110 SOFTWARE: PatentIn Ver. 2.0
111 SEQ ID NO 73
112 LENGTH: 680
113 TYPE: DNA
114 ORGANISM: Homo sapiens
115 FEATURE:
116 NAME/KEY: SITE
117 LOCATION: (7)
118 OTHER INFORMATION: n equals a,t,g, or c
119 FEATURE:
120 NAME/KEY: SITE
121 LOCATION: (9)
122 OTHER INFORMATION: n equals a,t,g, or c
123 FEATURE:
124 NAME/KEY: SITE
125 LOCATION: (15)
126 OTHER INFORMATION: n equals a,t,g, or c
127 FEATURE:
128 NAME/KEY: SITE
129 LOCATION: (16)
130 OTHER INFORMATION: n equals a,t,g, or c
131 FEATURE:
132 NAME/KEY: SITE
133 LOCATION: (22)
134 OTHER INFORMATION: n equals a,t,g, or c
135 PCT-US99-17130-73

alignment_scores:
Quality: 666.00 Length: 121
Ratio: 5.504 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-179 x PCT-US99-17130-73 ..
Align seg 1/1 to: PCT-US99-17130-73 from: 1 to: 680
1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLe 17
133 ATGATGCTACCCCAATGGCTGCTGCTGCTTCTCTCTCTCTCTCT 182
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34
183 CTTCCTCCTCACCAGGGGCTCACTTCTCCACAAATACAACTTTTGG 232
34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
233 AGCTCAAGGAGTCTTCATCCGGAACCGAGCTGCGAGACTGCGTCTGC 282
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 67
283 CAACGTGCTCCAGACAAATGCGAGTGCACACTGCGGGAAGGGGTCGA 332
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
333 GGGCAGTCTGTGTCACACGACGAGTGTCTTTGGCCAGATATAGCGTGT 382
84 roCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
383 CTGCGCTCGGAACCTGACTTGTATATATCAAAAGAAATGAGAAATGG 432
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 117
433 AGCATCGCCTATGGCGGTTGTGAGAAATTTGGAAGGACGAAAGTTGGCT 482

117 sLysMetPhePhe 121
118 |
483 GAAATGTTCTTC 495
seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-489-847-73
seq_documentation_block:
; Sequence 73, Application US/09489847
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-73
alignment_scores:
Quality: 666.00 Length: 121
Ratio: 5.504 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-179 x US-09-489-847-73 ..
Align seg 1/1 to: US-09-489-847-73 from: 1 to: 680
1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLe 17
133 ATGATGCTACCCCAATGGCTGCTGCTGCTTCTCTCTCTCTCTCT 182
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34
183 CTTCCTCCTCACCAGGGGCTCACTTCTCCACAAATACAACTTTTGG 232

us-09-215-435-179.rnmp

Thu Nov 15 10:52:01 2001

```

|||||
149 CTTCTCCTCCACACAGGGGCTCACTTTCTTCCACAAATAACACCTTTGG 198
|||||
34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
199 AGTCAAGAGAGTCTTGCATCCGGAACACGAGCTGCGAGACTGGCTGCTGC 248
|||||
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
249 CAACGTGCTCCAGACAAATTGCGAGTGCACCTGCGCGGAGAGGGGTCCGA 298
|||||
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
299 GGGCAGTCTGTGTCAAACGAGGAGTCTTTGGCCARTATAGAGCGTGC 348
|||||
84 roCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
|||||
349 CCTGCCTCGGGAACCTGACTTGTATATATCAAGAATGAGAAATGGCTT 398
|||||
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaLy 117
|||||
399 AGCATCGCCTATGGCGTGTGTCAGAAAATTGGAAGCGAGAGTTGGCTAA 448
|||||
117 sLysMetPhePhe 121
|||||
449 RAAATGTTCTTC 461

```

seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-4817

```

seq_documentation_block:
; Sequence 4817, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..461
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.80000019073486
; OTHER INFORMATION: seq FLLFFFLFLTRG/SL
US-60-197-873-4817

```

```

alignment_scores:
Quality: 661.00 Length: 121
Ratio: 5.508 Gaps: 0
Percent Similarity: 99.174 Percent Identity: 99.174

```

alignment_block:

```

US-09-215-435-179 x US-60-197-873-4817
Align seg 1/1 to: US-60-197-873-4817 from: 1 to: 533

```

```

1 MetMetLeuProGlnTrpLeuLeuLeuPhePhePheLe 17
99 ATGATGCTACCCCAATGGCTGTGCTGCTCTCTCTTCTTCT 148
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34

```

```

34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
233 AGTCAAGAGAGTCTTGCATCCGGAACACGAGCTGCGAGACTGGCTGCTGC 282
|||||
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
283 CAACGTGCTCCAGACAAATTGCGAGTGCACCTGCGCGGAGAGGGGTCCGA 332
|||||
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
333 GGGCAGTCTGTGTCAAACGAGGAGTCTTTGGCCARTATAGAGCGTGC 382
|||||
84 roCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
|||||
383 CCTGCCTCGGGAACCTGACTTGTATATATCAAGAATGAGAAATGGCTT 432
|||||
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaLy 117
|||||
433 AGCATCGCCTATGGCGTGTGTCAGAAAATTGGAAGCGAGAGTTGGCTAA 482
|||||
117 sLysMetPhePhe 121
|||||
483 GAAATGTTCTTC 495

```

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-4817

```

seq_documentation_block:
; Sequence 4817, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..461
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.80000019073486
; OTHER INFORMATION: seq FLLFFFLFLTRG/SL
US-09-834-366-4817

```

```

alignment_scores:
Quality: 661.00 Length: 121
Ratio: 5.508 Gaps: 0
Percent Similarity: 99.174 Percent Identity: 99.174

```

alignment_block:

```

US-09-215-435-179 x US-09-834-366-4817
Align seg 1/1 to: US-09-834-366-4817 from: 1 to: 533

```

```

1 MetMetLeuProGlnTrpLeuLeuLeuPhePhePheLe 17
99 ATGATGCTACCCCAATGGCTGTGCTGCTCTCTCTTCTTCT 148
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34

```

```
|||||
149 CTTCTCTCCACAGGGCTCACTTTCTCAACAAATACAACTTTGG 198
|||||
34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
199 AGCTCAAGAGTCTTCATCCGACAGGACTGCGAGACTGGCTGCTGC 248
|||||
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
249 CAAGGTGCTCCAGACAAATGCGAGTGCAGCTGCGGAGAGGGTCCGA 298
|||||
67 uGlySerLeuGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
299 GGGCAGTCTGCTCAACGACAGTGTCTTTGGCCARTATAGAGCGTGC 348
|||||
84 rOCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
|||||
349 CCGCTCGCGGACCTGACTGTATATATCAAGAATGAGAAATGGCTT 398
|||||
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 117
|||||
399 AGCATGCCCTATGGCGTGTCTCAGAAATTTGGAAGCAGCAAGTTGCTAA 448
|||||
117 slysMetPhePhe 121
|||||
449 RAAATGTCTTTC 461
```

```
seq_name: /cgnl_7/ptodata/1/pna/US089_COMB.seq:us-08-905-279-44

seq_documentation_block:
; Sequence 44, Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; TITLE OF INVENTION: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905.279
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 62..130
; IDENTIFICATION METHOD: Von Heijne matrix
```

```
;
; OTHER INFORMATION: score 9.8
; OTHER INFORMATION: seq FLLFFLLTRG/SL
us-08-905-279-44

alignment_scores:
Quality: 619.00 Length: 118
Ratio: 5.383 Gaps: 0
Percent Similarity: 97.458 Percent Identity: 95.763

alignment_block:
US-09-215-435-179 x US-08-905-279-44 ..
Align seg 1/1 to: US-08-905-279-44 from: 1 to: 421

1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLe 17
|||||
62 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTTCTCTCTTTCT 111
|||||
17 uPheLeuLeuThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuG 34
|||||
112 CTTCTCTCCACAGGGCTCACTTTCTCAACAAATACAACTTTGG 161
|||||
34 luLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCys 50
|||||
162 AGCTCAAGGAGKSKGCATSGGGAACCCAGGACTGCGAGACTGGCTGCTGC 211
|||||
51 GlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG1 67
|||||
212 CAACGTGCTCCAGACAAATGCGAGTGCAGCTGCGGAGAGGGTCCGA 261
|||||
67 uGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysP 84
|||||
262 GGGCAGTCTGTCTCAACGACAGTGTCTTTGGCCARTATAGAGCGTGC 311
|||||
84 rOCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 100
|||||
312 CCTGCTCGGGAACCTGACTTGTATATATCAAGAATGAGAAATGGCTT 361
|||||
101 SerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 117
|||||
362 AGCATGCCCTATGGCGTGTCTCAGAAATTTGGAAGCAGCAAGTTGCTAG 411
|||||
117 slys 118
|||
412 RAAA 415

seq_name: /cgnl_7/ptodata/1/pna/US081_COMB.seq:us-08-196-481-520

seq_documentation_block:
; Sequence 520, Application US/08196481
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William
; APPLICANT: Rosen, Craig
; APPLICANT: Ruben, Steve
; APPLICANT: Dillon, Patrick
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences, and Expression
; NUMBER OF SEQUENCES: 1719
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
; ADDRESSEE: Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
```

Products

```
seq_name: /cgml_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481B-520
seq_documentation_block:
; Sequence 520, Application US/08196481B
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 7
; FILE REFERENCE: P007
; CURRENT APPLICATION NUMBER: US/08/196.481B
; CURRENT FILING DATE: 1994-02-15
; NUMBER OF SEQ ID NOS: 1721
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 520
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (220)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (317)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (354)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (356)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (366)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (378)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (425)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (434)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (437)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (440)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (449)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (457)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (458)
```

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (463)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (482)
; OTHER INFORMATION: n equals a,t,g, or c
; US-08-196-481B-520

alignment_scores:
Quality: 602.00 Length: 112
Ratio: 5.423 Gaps: 0
Percent Similarity: 99.107 Percent Identity: 97.321

alignment_block:
US-09-215-435-179 x US-08-196-481B-520 ..
Align seg 1/1 to: US-08-196-481B-520 from: 1 to: 488
10 LeuPheLeuPhePhePheLeuPheLeuThrArgGlySerLeuSe 26
|||||
14 CAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63
|||||
26 rProThrLysTyrAsnLeuLeuLeuLysGluSerCysIleArgAsnG 43
|||||
64 TCCACAAAAATACAACTTTTGAGAGTCAAGGAGTCTTGCATCCGGAACC 113
|||||
43 InAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGluSer 59
|||||
114 AGGACTGCGAGACTGGCTGCTGCCAACGTCTCCAGACAAATTGCGAGTCG 163
|||||
60 HisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnValPh 76
|||||
164 CACTGCGCGAGAGGGTCCGAGGCGAGTCTGTGTCAACACGCGGTGTT 213
|||||
76 ePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysIleT 93
|||||
214 CTTTGGNCAATATAGAGCGTGTCCCTGCCTGCGGAACCTGACTTGTATAT 263
|||||
93 YrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgCysGlnLys 109
|||||
264 ATTCAAAGAATGAGAAATGGCTTAGCATCGCCTATGCCCTTGTCTCAGAA 313
|||||
110 IleGlyArgGlnLysLeuAlaLysLysMetPhePhe 121
|||||
314 ATTNGAAGGCAGAGTTGGTTAAGGAATGTTCTTT 349
|||||

seq_name: /cgml_7/ptodata/1/pna/US6018_COMB.seq:US-60-184-698-48
seq_documentation_block:
; Sequence 48, Application US/60184698
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M. Jones, Anissa L.
; APPLICANT: Lincoln, Stephen E. Yu, Jimmy Y.
; APPLICANT: Russo, Frank D. Greenawalt, Lila B.
; APPLICANT: Spiro, Peter A. Panzer, Scott R.
; APPLICANT: Banville, Steve C. Roseberry, Ann M.
; APPLICANT: Bratcher, Shawn R. Wright, Rachel J.
; APPLICANT: Dufour, Gerard E. Chen, Wensheng
; APPLICANT: Cohen, Howard J. Liu, Tommy
; APPLICANT: Rosen, Bruce Yap, Pierre E.
; APPLICANT: Shah, Purvi Amshey, Stefan
; APPLICANT: Chalup, Michael S. Fong, Willy Tuen
; APPLICANT: Hillman, Jennifer L. BIOCHEMICAL PATHWAY MOLECULES
; TITLE OF INVENTION: PT-0115 P
; FILE REFERENCE: PT-0115 P
; CURRENT APPLICATION NUMBER: US/60/184,698
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 1132
; SOFTWARE: PERL Program
; SEQ ID NO 48
; LENGTH: 463
```

us-09-215-435-179.rnmp

Thu Nov 15 10:52:01 2001

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 474641.1
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 474641.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 347-368
; OTHER INFORMATION: a, t, c, g, or other
US-60-184-698-48

alignment_scores:
  Quality: 497.00      Length: 88
  Ratio: 5.648        Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-215-435-179 x US-60-184-698-48 ..
  Align seg 1/1 to: US-60-184-698-48 from: 1 to: 463
    34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
    |||||
    61 GAGCTCAAGGAGTCTTCATCCGGAACCGAGACTGCGAGACTGGCTGCTG 110
    |||||
    50 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 67
    |||||
    111 CCAAGTCTCCAGACAAATTCGAGTGCACACTGCGGAGAGGGTCCG 160
    |||||
    67 LuGlySerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCys 83
    |||||
    161 AGGGCAGTCTGTGTCACACGAGGTGTTCTTTGGCCAATATAGAGCGTGT 210
    |||||
    84 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 100
    |||||
    211 CCTGCTCGGGAACCTGACCTGTATATATTCAGAAATGGAAGATGGCT 260
    |||||
    100 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 117
    |||||
    261 TAGCATCGCTATGCGCGTTGTACAGAAATTTGGAAGGAGGAGTTGGCTA 310
    |||||
    117 ySLysMetPhePhe 121
    |||||
    311 AGAAATGTTCTTC 324

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-184-698-694

seq_documentation_block:
; Sequence 694, Application US/60184698
; GENERAL INFORMATION:
; APPLICANT: Jones, Anissa L.
; TITLE OF INVENTION: Yu, Jimmy Y.
; FILE REFERENCE: Greenawalt, Lila B.
; CURRENT APPLICATION NUMBER: Panzer, Scott R.
; PRIOR FILING DATE: Roseberry, Ann M.
; NUMBER OF SEQ ID NOS: Wright, Rachel J.
; SOFTWARE: Chen, Wensheng
; FASTSEQ for Windows Version 3.0
; SEQ ID NO 1: Liu, Tommy
; TYPE: yap, Pierre E.
; ORGANISM: Amsheer, Stefan
; NAME/KEY: Fong, Willy Tuen
; LOCATION: PT-0115 P
; CURRENT APPLICATION NUMBER: BIOCHEMICAL PATHWAY MOLECULES
; CURRENT FILING DATE: US/60/184,698
; NUMBER OF SEQ ID NOS: 1132
; SOFTWARE: PERL Program
; SEQ ID NO 694
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens

```



```

alignment_block:
US-09-215-435-179 x PCT-US00-31337-1 ..
Align seg 1/1 to: PCT-US00-31337-1 from: 1 to: 523

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
|||||
91 GAGCTCAAGGAGTCTTGCATCCGGAACAGGACTCGGAGACTGGCTGCTG 140

50 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 67
|||||
141 CCAACGTCCTCCAGACAAATTCGAGTCGCACCTGCGCGGAGAGGGTCCG 190

67 luGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 83
|||||
191 AGGGCAGTCTGTCTCAACGCGAGGTGTCTTTGGCCATATAGACGGTGT 240

84 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 100
|||||
241 CCTCGCTCGCGAACCTGACTGTATATATTCAAAGAATGAGAAATGGCT 290

100 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 117
|||||
291 TAGCATCGCTATGCGCGTGTGTGAGAAAAATTCGAGGAGGAGAGTTGGCTA 340

117 ysLysMetPhePhe 121
|||||
341 AGAAAATGTTCTTC 354

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-443-634-1

seq_documentation_block:
; Sequence 1, Application US/09443634
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Colipase-Like Protein - 1
; FILE REFERENCE: 99-85
; CURRENT APPLICATION NUMBER: US/09/443.634
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(354)
US-09-443-634-1

alignment_scores:
Quality: 497.00 Length: 88
Ratio: 5.648 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179 x US-09-443-634-1 ..
Align seg 1/1 to: US-09-443-634-1 from: 1 to: 523

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
|||||
91 GAGCTCAAGGAGTCTTGCATCCGGAACAGGACTCGGAGACTGGCTGCTG 140

50 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 67
|||||
141 CCAACGTCCTCCAGACAAATTCGAGTCGCACCTGCGCGGAGAGGGTCCG 190

67 luGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 83
|||||
191 AGGGCAGTCTGTCTCAACGCGAGGTGTCTTTGGCCATATAGACGGTGT 240

```

```

84 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 100
|||||
241 CCTCGCTCGCGAACCTGACTGTATATATTCAAAGAATGAGAAATGGCT 290

100 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 117
|||||
291 TAGCATCGCTATGCGCGTGTGTGAGAAAAATTCGAGGAGGAGAGTTGGCTA 340

117 ysLysMetPhePhe 121
|||||
341 AGAAAATGTTCTTC 354

seq_name: /cgnl_7/ptodata/1/pna/US097A_COMB.seq:US-09-713-067-1

seq_documentation_block:
; Sequence 1, Application US/09713067
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: EDUCATIONAL KIT AND METHOD USING
; FILE REFERENCE: COLIPASE-LIKE PROTEIN-1
; FILE REFERENCE: 99-85US
; CURRENT APPLICATION NUMBER: US/09/713.067
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/166,626
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(354)
US-09-713-067-1

alignment_scores:
Quality: 497.00 Length: 88
Ratio: 5.648 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179 x US-09-713-067-1 ..
Align seg 1/1 to: US-09-713-067-1 from: 1 to: 523

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
|||||
91 GAGCTCAAGGAGTCTTGCATCCGGAACAGGACTCGGAGACTGGCTGCTG 140

50 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 67
|||||
141 CCAACGTCCTCCAGACAAATTCGAGTCGCACCTGCGCGGAGAGGGTCCG 190

67 luGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 83
|||||
191 AGGGCAGTCTGTCTCAACGCGAGGTGTCTTTGGCCATATAGACGGTGT 240

84 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 100
|||||
241 CCTCGCTCGCGAACCTGACTGTATATATTCAAAGAATGAGAAATGGCT 290

100 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 117
|||||
291 TAGCATCGCTATGCGCGTGTGTGAGAAAAATTCGAGGAGGAGAGTTGGCTA 340

117 ysLysMetPhePhe 121
|||||
341 AGAAAATGTTCTTC 354

```

us-09-215-435-179.rnp

Thu Nov 15 10:52:01 2001

Thu Nov 15 10:52:01 2001

```

seq_documentation_block:
; Sequence 9, Application US/09703559
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Insulin-Responsive Sequence DNA Binding Protein-1 and Methods to
; FILE REFERENCE: E056 1010
; CURRENT APPLICATION NUMBER: US/09/703.559
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 3087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3084)
; US-09-703-559-9

alignment_scores:
Quality: 74.50 Length: 63
Ratio: 2.328 Gaps: 5
Percent Similarity: 50.794 Percent Identity: 33.333

alignment_block:
US-09-215-435-179 x US-09-703-559-9 ..
Align seg 1/1 to: US-09-703-559-9 from: 1 to: 3087

44 AspCysGluThrGly.....CysCysGlnArgAlaProAs 55
:::|||||
205 GAGTCCCCCGAAGGCTTCATGGCGCTGGAGTGCAGGGAGAGAGTCCCGA 254
:::|||||
55 pAsnCysGlu.....SerHisCysAlaGluLysGlySerGluG 68
:::|||||
255 TGACTGTGAGTGGCGCAACGAGGAGGAGATGCTG.....GGCGCCAACA 298
:::|||||
68 lYSerLeuCysGlnThrGlnVal.....PhePheGly..... 78
:::|||||
299 CCACCTCTGCAGTGGCGCCCTGCAACATGAACACACAGTGC 348
:::|||||
79 GlnTyrArgAlaCysProCysLeuArgAsnLeuThrCys 91
:::|||||
349 GAATCAGACGCCATGCCCTGCAACATGAACACACAGTGC 387
:::|||||

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-703-559-7

seq_documentation_block:
; Sequence 7, Application US/09703559
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Insulin-Responsive Sequence DNA Binding Protein-1 and Methods to
; FILE REFERENCE: E056 1010
; CURRENT APPLICATION NUMBER: US/09/703.559
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4524
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(58)
; NAME/KEY: exon
; LOCATION: (59)..(3082)
; NAME/KEY: 3'UTR
; LOCATION: (3083)..(4524)
; US-09-703-559-7

```

```

alignment_scores:
Quality: 74.50 Length: 63
Ratio: 2.328 Gaps: 5
Percent Similarity: 50.794 Percent Identity: 33.333

alignment_block:
US-09-215-435-179 x US-09-703-559-7 ..
Align seg 1/1 to: US-09-703-559-7 from: 1 to: 4524

44 AspCysGluThrGly.....CysCysGlnArgAlaProAs 55
:::|||||
263 GAGTCCCCCGAAGGCTTCATGGCGCTGGAGTGCAGGGAGAGAGTCCCGA 312
:::|||||
55 pAsnCysGlu.....SerHisCysAlaGluLysGlySerGluG 68
:::|||||
313 TGACTGTGAGTGGCGCAACGAGGAGGAGATGCTG.....GGCGCCAACA 356
:::|||||
68 lYSerLeuCysGlnThrGlnVal.....PhePheGly..... 78
:::|||||
357 CCACCTCTGCAGTGGCGCCCTGCAACATGAACACACAGTGC 406
:::|||||
79 GlnTyrArgAlaCysProCysLeuArgAsnLeuThrCys 91
:::|||||
407 GAATCAGACGCCATGCCCTGCAACATGAACACACAGTGC 445
:::|||||

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-970-966-206

seq_documentation_block:
; Sequence 206, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Moles, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jlangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 2364
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-970-966-206

alignment_scores:
Quality: 74.00 Length: 103
Ratio: 1.321 Gaps: 4
Percent Similarity: 54.369 Percent Identity: 20.388

alignment_block:
US-09-215-435-179 x US-09-970-966-206 ..
Align seg 1/1 to: US-09-970-966-206 from: 1 to: 2364

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
||| |||
1567 GAGGAACGGAGAGTGCACGGTCAACGAGGAGTGTCTCTCCACGACGTG 1616
||| |||
50 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAla. 62
||| |||
1617 CCTGATGACCGAGTGGGGCCGAGTGGGACGAGTGCAGCGCCACCTCGGGCA 1666
||| |||
63 .....GluTysGlySerGluGly 68
||| |||
1667 TGGGCATGAAGAACGCGCACCGCATGATCAAGATGAACCCCGCAGATGTC 1716
||| |||

```

```
69 SerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1717 TCCATGTGCAAGCCGAGACATCAGCAGAGAGTGCATGATGCCAGA 1766

85 SerArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 100
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1767 GTGCCACACCATCCCATCTGCTGCTGCCATGGTCCGAGTGGAGTGACT 1816

100 euSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1817 GCAGCGTGACCTGCGGGAAGGCATGCCGAACC...CGACAGCGGATGCTC 1863

117 LysLysMet 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1864 AAGTCTCTG 1872

seq_name: /cgnl_7/ptodata/1/pna/us09_NEW_COMB.seq:us-09-970-966-184
seq_documentation_block:
; Sequence 184, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 3079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-970-966-184
```

```
alignment_scores:
  Quality: 74.00      Length: 103
  Ratio: 1.321       Gaps: 4
  Percent Similarity: 54.369  Percent Identity: 20.388

alignment_block:
US-09-215-435-179 x US-09-970-966-184
..
Align seg 1/1 to: US-09-970-966-184 from: 1 to: 3079

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
|||
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1651 GAGGAATGGAGAGTGCACGGTCAACGAGGAGTGCTCTCCAGCAGCTG 1700

50 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAla. 62
|
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1701 CCTGATGACCGAGTGGGCGAGTGGGACGAGTGCAGCGCCACTCGGCA 1750

63 .....GluLysGlySerGluGly 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1751 TGGGCATGAAGAACGCCCATCATGATCAAGATGAACCCCGCATGGC 1800

69 SerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1801 TCCATGTGCAAGCCGAGACATCAGCAGAGAGTGCATGATGCCAGA 1850

85 sLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 100
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1851 GTGCCACACCATCCCATCTGCTGCTGCCATGGTCCGAGTGGAGTGACT 1900

100 euSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1901 GCAGCGTGACCTGCGGGAAGGCATGCCGAACC...CGACAGCGGATGCTC 1947
```

```
117 LysLysMet 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1948 AAGTCTCTG 1956

seq_name: /cgnl_7/ptodata/1/pna/us09_NEW_COMB.seq:us-09-970-966-177
seq_documentation_block:
; Sequence 177, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 177
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-970-966-177

alignment_scores:
  Quality: 74.00      Length: 103
  Ratio: 1.321       Gaps: 4
  Percent Similarity: 54.369  Percent Identity: 20.388

alignment_block:
US-09-215-435-179 x US-09-970-966-177
..
Align seg 1/1 to: US-09-970-966-177 from: 1 to: 3999

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 50
|||
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1078 GAGGAACGGAGAGTGCACGGTCAACGAGGAGTGCTCTCCAGCAGCTG 1127

50 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAla. 62
|
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1128 CCTGATGACCGAGTGGGCGAGTGGGACGAGTGCAGCGCCACTCGGCA 1177

63 .....GluLysGlySerGluGly 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1178 TGGGCATGAAGAACGCCCATCATGATCAAGATGAACCCCGCATGGC 1227

69 SerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1228 TCCATGTGCAAGCCGAGACATCAGCAGAGAGTGCATGATGCCAGA 1277

85 sLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 100
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1278 GTGCCACACCATCCCATCTGCTGCTGCCATGGTCCGAGTGGAGTGACT 1327

100 euSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1328 GCAGCGTGACCTGCGGGAAGGCATGCCGAACC...CGACAGCGGATGCTC 1374

117 LysLysMet 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1375 AAGTCTCTG 1383

seq_name: /cgnl_7/ptodata/1/pna/us09_NEW_COMB.seq:us-09-976-594-353
seq_documentation_block:
; Sequence 353, Application US/09976594
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROI
```

Thu Nov 15 10:52:01 2001

us-09-215-435-179.rnnpn

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-28

alignment_scores:
  Quality: 72.50      Length: 133
  Ratio: 1.250        Gaps: 6
  Percent Similarity: 43.609  Percent Identity: 24.812

alignment_block:
  US-09-215-435-179 x US-09-886-055-28/rev ..
  Align seg 1/1 to reverse of: US-09-886-055-28 from: 1 to: 1026

21 ThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysG1 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
798 ACCGGTGGAGGCTGCTCCGCGACTGCCCTCCAGCGCGCATGTGGATCA 749

37 user.....CysIleArgAsnGlnAspCys...GluThrGlyC 49
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
748 CAGCCCGAGCGATGAACCCATTAAGACACCGCTGATAGCTAACACCGTGGCT 699

49 ysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGly 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 GTCGTGGATGAAGCCACACATGGCGAAGAGAAGGAGTTCATTGAGACTGGI 649

65 .....
66 .....S 66
648 GTCACCTGCAGAGATGCCACGACGTCAGCGGATATCCGAGAAGAAGCTAT 599

66 .....
598 TCATCTTCGGGAGCGGAGAGCTCAGCGGAGAGGTGAGGGTTGTGTGA 549

66 erGluGlySerLeuCysGlnThrGlnValPheGlyGlnTyrArgAla 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
548 ACAAGGACACTCACTGCCCCCAGCGGCTGATGCTCCACCAAGGCCAG 499

83 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 GCATAGACGCTCGACATAGCTGTTGTATA.....GAGAAGTG 461

99 pleuSer.....IleAlaTyrGlyArgCysGlnLys 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
460 GGTTCCTGATGGCCACGTCAGCGGTCATAGGCCATGGC...TCCACGCAA 415

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-189-455

seq_documentation_block:
; Sequence 455, Application US/09978189
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
```

APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C7
CURRENT APPLICATION NUMBER: US/09/978,189
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640

```
seq_name: /cgml_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-192-455

seq_documentation_block:
; Sequence 455, Application US/09978192
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
```


1	PRIOR APPLICATION NUMBER: 60/083555	
2	PRIOR FILING DATE: 1998-04-29	
3	PRIOR APPLICATION NUMBER: 60/083559	
4	PRIOR FILING DATE: 1998-04-29	
5	PRIOR APPLICATION NUMBER: 60/083500	
6	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083742	
8	PRIOR FILING DATE: 1998-04-30	
9	PRIOR APPLICATION NUMBER: 60/084366	
10	PRIOR FILING DATE: 1998-05-05	
11	PRIOR APPLICATION NUMBER: 60/084414	
12	PRIOR FILING DATE: 1998-05-06	
13	PRIOR APPLICATION NUMBER: 60/084441	
14	PRIOR FILING DATE: 1998-05-06	
15	PRIOR APPLICATION NUMBER: 60/084637	
16	PRIOR FILING DATE: 1998-05-07	
17	PRIOR APPLICATION NUMBER: 60/084639	
18	PRIOR FILING DATE: 1998-05-07	
19	PRIOR APPLICATION NUMBER: 60/084640	
20	PRIOR FILING DATE: 1998-05-07	
21	PRIOR APPLICATION NUMBER: 60/084598	
22	PRIOR FILING DATE: 1998-05-07	
23	PRIOR APPLICATION NUMBER: 60/084600	
24	PRIOR FILING DATE: 1998-5-07	
25	PRIOR APPLICATION NUMBER: 60/084627	
26	PRIOR FILING DATE: 1998-05-07	
27	PRIOR APPLICATION NUMBER: 60/084643	
28	PRIOR FILING DATE: 1998-05-07	
29	PRIOR APPLICATION NUMBER: 60/085339	
30	PRIOR FILING DATE: 1998-05-13	
31	PRIOR APPLICATION NUMBER: 60/085338	
32	PRIOR FILING DATE: 1998-05-13	
33	PRIOR APPLICATION NUMBER: 60/085323	
34	PRIOR FILING DATE: 1998-05-13	
35	PRIOR APPLICATION NUMBER: 60/085582	
36	PRIOR FILING DATE: 1998-05-15	
37	PRIOR APPLICATION NUMBER: 60/085700	
38	PRIOR FILING DATE: 1998-05-15	
39	PRIOR APPLICATION NUMBER: 60/085689	
40	PRIOR FILING DATE: 1998-05-15	
41	PRIOR APPLICATION NUMBER: 60/085579	
42	PRIOR FILING DATE: 1998-05-15	
43	PRIOR APPLICATION NUMBER: 60/085580	
44	PRIOR FILING DATE: 1998-05-15	
45	PRIOR APPLICATION NUMBER: 60/085573	
46	PRIOR FILING DATE: 1998-05-15	
47	PRIOR APPLICATION NUMBER: 60/085704	
48	PRIOR FILING DATE: 1998-05-15	
49	PRIOR APPLICATION NUMBER: 60/085697	
50	PRIOR FILING DATE: 1998-05-15	

```

alignment_scores:      72.50      Length: 86
                       Ratio:  1.726      Gaps: 4
                       Percent Similarity: 48.837      Percent Identity: 26.744

alignment_block:
US-09-215-435-179 x US-09-978-192-455      ..

Align seg 1/1 to: US-09-978-192-455      from: 1 to: 1518

22 ArgGlySerLeuSerProThrLysTyrAsnLeuLeuLysGluSe 38
||| :||| ||| :|||
648 AGAACCACTTGTCTTCAAAATGTATCACACCAAGGACAGAAGGTC 697

38 r...CysIleArgAsnGlnAspCysGluThrGlyCysGlnArgAlap 54
||| :||| :||| :||| :|||
698 TGTTTGTCTCGGGTCATCAGACTGCCTCAGGATGTTGT..... 737

54 roAspAsnCysGluSerHis.....CysAlaGluLysGly 65
||| :|||
738 .....TGTGCTAGACACTTCTGGTCCCAAGATCTGTAACCTGTCCTG 779

```


Thu Nov 15 10:52:01 2001

[illegible]

```

alignment_scores:
  Quality: 72.50      Length: 86
  Ratio: 1.726       Gaps: 4
  Percent Similarity: 48.837   Percent Identity: 26.744

alignment_block:
US-09-215-435-179 x US-09-974-298-88  ..

Align seg 1/1  to: US-09-974-298-88  from: 1  to: 1797

22  ArgGlySerLeuSerProThrLysTyrAsnLeuLeuLeuLeuLysGluSe 38
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:
662 AGAACCACTGTCTTCAAAAATGATACACACCAAGGACCAAGAGGTC 711
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

38  r...CysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaP 54
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:
712 TGTTCCTCCGGTCATCAGACTGTGCTCAGGATTGTGTT 751
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

54  roAspAsnCysGluSerHis.....CysAlaGluLysGly 65
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:
752 .....TGTGCTAGACACTCTGTGGTCCAAGATCTGTAAACCTGTCCTG 793
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

66  SerGluGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 79
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

794 AAGAAGGTCAGTGTGTACCAACATAGGAGAAAGGCTCATGGACT 843
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

80  .....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrS 94
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:
844 AGAAATATCCAGCGTTGTACTGTGGAGAAGGTCTGTTCGCCGGATAC 893
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

94  erLysAsn 96
  :|||:|||||:

894 AGAAAGAT 901

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:us-09-976-594-594

seq_documentation_block:
; Sequence 1085, Application US/09976594
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURE
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976, 594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1085
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 2481150CBI
US-09-976-594-1085

alignment_scores:
  Quality: 72.50      Length: 86
  Ratio: 1.726       Gaps: 4
  Percent Similarity: 48.837   Percent Identity: 26.744

alignment_block:
US-09-215-435-179 x US-09-976-594-1085  ..

Align seg 1/1  to: US-09-976-594-1085  from: 1  to: 1797

22  ArgGlySerLeuSerProThrLysTyrAsnLeuLeuLeuLysGluSe 38
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:
662 AGAACCACTGTCTTCAAAAATGATACACACCAAGGACCAAGAGGTC 711
  ||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||:

```

Thu Nov 15 10:52:01 2001

```

45 CysGluThrGlyCysGlnArgAlaProAspAsnCysGluSerHisCy 61
|||:||||| :||| :||| :||| :||| :||| :||| :|||
19838 ATAGAAAGGGGAAACGCTCAA.....CGTGCTCTTCCAAAATCATAC 19881
61 sAlaGluLysGlySerGluGlySerLeuLeuCysGlnThrGlnValPhePheG 78
:|||||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
19882 CACCGAGAAAGAGATATCGGTCAG.....TTTACCG 19913
78 lYgIntyrArg..... 81
|||:|||||
19914 GTCGTTATCGCTCTCCCTGTTTCAGGTTATCCCTCATCGTAAATGCGGT 19963
82 ...AlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnG1 97
|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
19964 CAGATGTCGCGG...ATCAGAAAGGGTAGAGCGTAAATCGTTAAATGGA 20010
97 uLysTrpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArg.... 112
:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
20011 TATTTGGCTGATGATCACC.....GAATGTGAGAAAAATGTCGCTTTG 20054
113 .....GlnLysLeuAlaLys 117
|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
20055 TCGGTTTAGCCACAAACTCACTANA 20080

```

```

38 r...CysIleArgAsnGlnAspCysGluThrGlyCysGlnArgAlap 54
| :|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
712 TGTGTGTCTCCGCTCATCAGACTGTCCTCAGGATTGTG..... 751
54 roAspAsnCysGluSerHis.....CysAlaGluLysGly 65
|||:|||||
752 .....TGTCTAGACACTCTCTGCTCAAGATCTGTAAACCTGTCCTG 793
66 SerGluGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 79
:|||||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
794 AAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAGGCTCTCATGGACT 843
80 .....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrS 94
|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
844 AGAATATTCCAGCGTGTGTACTGTGGAGAGGTCGTCTTCCCGGATAC 893
94 erLysAsn 96
:|||||:
894 AGAAGAT 901

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-344

```

seq_documentation_block:
; Sequence 344, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2000-06-25, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 344
; LENGTH: 26092
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12281)..(13480)
; US-09-897-516-344

```

```

alignment_scores:
Quality: 72.50 Length: 142
Ratio: 1.098 Gaps: 9
Percent Similarity: 46.479 Percent Identity: 26.056

```

```

alignment_block:
US-09-215-435-179 x US-09-897-516-344 ..
Align seg 1/1 to: US-09-897-516-344 from: 1 to: 26092
5 GlnTrpLeuLeuLeu.....LeuPh 11
|||||:|||||
19709 CAGTGGTTACTCTTACCGGCACAGATAACTGCTGGTAAATATAATGCT 19758
11 eLeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProt 28
:|||||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
19759 TATAGCTTTCTACCAAGCGCTTATCTGACAGTCGTTAAGCTCCCCCCA 19808
28 hrLysTyrAsnLeuLeuGluLeuLysGluSerCysIleArgAsnGlnAsp 44
|||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
19809 CACAA.....ATGAGCGATGTCATATTTGCTG..... 19837

```


us-09-215-435-179.1st

Thu Nov 15 10:52:02 2001

```

423 TTTTGGAGCTCAAGGAGTCTTGATCCGACACAGGACTGCGAGACTGGC 374
49 CysCysGlnArgAlaProAsnGlnSerHis.CysAlaGluLysG 65
|||||
373 TCGTCCCAAGCGTCTCCAGACAATTCGAGTCGCGACTNGCCGCGAGAAGG 324
|||||
65 lYSerGluGlySerLeuGlnThrGlnValPhePheGlyGlnTyrArg 81
|||||
323 GGTCCAGGCGAGTCTGTCAACAGCAGGCTGCTTTGGCCATATAGA 274
|||||
82 AlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 98
|||||
273 GCGTGTCCCTGCTG.CGGAACCTGACTGTATATATCAAGAATGAGAA 225
|||||
98 sTrpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysL 115
|||||
224 ATGGCTTAGCATGCCCCATGCGCGTGTGCAGAAATTTGGAAGCGAGAAGT 175
|||||
115 euAlaLysLysMetPhePhe 121
|||||
174 TGGCTAAGAAATGTTCTTC 155

seq_name: gb_gss29:AZ575911

seq documentation block: 127 bp DNA GSS 06-DEC-2000
LOCUS AZ575911
DEFINITION AST-T22D0047 Genetrap T47D Human Breast Carcinoma Library Homo sapiens genomic 5', DNA sequence.
ACCESSION AZ575911
VERSION AZ575911.1 GI:11562222
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Henkel,G., Liyanage,M., Pratt,E., Huang,D., Riley,M., Bernardino,A., Durick,K. and Pollok,B.
TITLE Exon-trap tags from a T47D GenomeScreen(TM) Library
JOURNAL Unpublished (2000)
COMMENT Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com
Pools of cells were isolated from a GenomeScreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun cloned in pAMP-1 and used to transform DH5-alpha competent bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.
Class: exon-trapped.
Location/Qualifiers
1. .127
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/clone_line="T47D"
/note="Organ: Breast; Vector: pAMP-1; 3' RACE of total RNA from genetrap pools; shotgun clone in pAMP-1 and used to transform DH5-alpha competent bacteria."
BASE COUNT 28 a 34 c 43 g 21 t 1 others
ORIGIN

alignment_scores:
Quality: 196.00 Length: 40
Ratio: 5.026 Gaps: 0
Percent Similarity: 97.500 Percent Identity: 97.500

alignment_block:
US-09-215-435-179 x AZ575911 ..
Align seg 1/1 to: AZ575911 from: 1 to: 127

34 GluLeuLysGluSerCysIleArgAsnGlnAspCysGlnThrGlyCysCy 50
|||||
1 GAGCTCAAGGAGTCTTGTCATCCGGAACCAAGGACTGCGAGACTGGCTGCTG 50
|||||
50 sGlnArgAlaProAsnGlnSerHisCysAlaGluLysGlySerG 67
|||||
51 CCAACGTGCTCCAGACAA.TGGGAGTCGCACATGCGCGAGAGAGGGGTCCN 99
|||||
67 luGlySerLeuGlnThr 73
|||||
100 AGGCAGTCTGTGTCAACG 119

seq_name: gb_est88:BF509739

seq documentation block: 293 bp mRNA EST 06-DEC-2000
LOCUS BF509739
DEFINITION UI-H-B14-agp-b-05-0-UI.s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone IMAGE:3087296 3', mRNA sequence.
ACCESSION BF509739
VERSION BF509739.1 GI:11593037
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, ph.D.
Email: cgaabs@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=yes.
Location/Qualifiers
1. .293
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3087296"
/clone_lib="NCI_CGAP_Sub8"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub8 is a subtracted library derived from NCI_CGAP_Sub5. The NCI_CGAP_Sub8 library had 2.5 million recombinants. A single-stranded DNA preparation of NCI_CGAP_Sub5 was used as a tracer in a subtractive hybridization with a driver comprising: a pool of clones from NCI_CGAP_Sub5 (IMAGE clone ids 2732833-2737415, 3068040-3069191; 25% of the driver population), a pool of clones from NCI_CGAP_Sub4 (IMAGE clone ids 2723592-2729326; 25% of the driver population), NCI_CGAP_Sub6 (pool AIF-AJU, IMAGE ids 2728969-2733190; 25% of the driver population), and NCI_CGAP_Sub7 (IMAGE ids 3069192-3072238, 3081864-3084550; 25% of the driver population). Subtraction was performed as previously described [Bonaldo, Lennon &

```


Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.

TAG_LIB=NCI_CGAP_Br2
TAG_TISSUE=breast
TAG_SEQ=AAACC

BASE COUNT 63 a 75 c 72 g 83 t
ORIGIN

alignment_scores:
Quality: 164.50 Length: 38
Ratio: 5.306 Gaps: 1
Percent Similarity: 81.579 Percent Identity: 78.947

alignment_block:
US-09-215-435-179 x BF509739/rev ..

Align seg 1/1 to reverse of: BF509739 from: 1 to: 293

48 GlyCysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLy 64

284 GGCTGCTGCCAACGTCCTCCAGACAAATGGCAGTGCCTCGCGGAGAA 235

64 sGlySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrA 81

234 GGGGTCCGAGGCGACGTGTGTCAAAACGACGTG.....C 200

81 rgAlaCysProCys 85

199 CAGCTACTCCATGC 186

seq_name: gb_est69:BE047055

seq_documentation_block: 540 bp mRNA EST 08-JUN-2000
LOCUS BE047055
DEFINITION hg55909.x1 NCI_CGAP_Pan3 Homo sapiens cDNA clone IMAGE:3123328 3'
similar to gb:J02883 COLIPASE PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION BE047055

VERSION BE047055.1 GI:8364108

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 540)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -400P from Gibco

High quality sequence stop: 397.

Location/Qualifiers

1..540

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3123328"

/lab_host="NCI_CGAP_Pan3"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 1.1 kb. Library constructed by Life

Technologies."

BASE COUNT 123 a 131 c 161 g 124 t 1 others
ORIGIN

alignment_scores:
Quality: 143.50 Length: 113
Ratio: 2.080 Gaps: 3
Percent Similarity: 61.062 Percent Identity: 30.973

alignment_block:

US-09-215-435-179 x BE047055/rev ..

Align seg 1/1 to reverse of: BE047055 from: 1 to: 540

7 LeuLeuLeuLeuPheLeuPhePhePheLeuPheLeuLeuThrArgG1 23

500 CTCCTGCTTGCCTTTTGTGGCCTATGAGGCTCTGGCCCCCGGGG 451

23 yserLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysI 40

450 GATCATTATTAAC.....TTGGAGACCGGTGAGCTCTGCA 416

40 leArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsn 56

415 TCAATAGTGCCTAGTGTAAAGACAAATGCTGCCAGCATTCRAAGTCGCTG 366

57 CysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnTh 73

365 GCCTTGGCCCGCTGCACATCCATGCCAGCGAGAACAGCGAGTGCTCTGT 316

73 rGlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeu 90

315 CAAGACGCTTATGGGATTACTACAAGTGTCCCTGTGAGCGTGGCCTCA 266

90 hrCysIleTyrSerLysAsnGluLysTrpLeu.....Ser 101

265 CCTGT.....GAGGAGACAAAGACCATTTGGGCTCCATCACCAC 225

102 lleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLys 114

224 ACCAACTTGGCATTTGCCATGACGNTGGACGCTCCCAAG 186

seq_name: gb_est69:BE047063

seq_documentation_block: 529 bp mRNA EST 08-JUN-2000

LOCUS BE047063

DEFINITION hg55907.x1 NCI_CGAP_Pan3 Homo sapiens cDNA clone IMAGE:3123325 3'

similar to gb:J02883 COLIPASE PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION BE047063

VERSION BE047063.1 GI:8364116

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 529)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Possible reversed clone: polyT not found

Seq primer: -400P from Gibco

High quality sequence stop: 412.

Thu Nov 15 10:52:02 2001

us-09-215-435-179.rst

```

FEATURES
  source
    Location/Qualifiers
      1..529
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:312325"
      /clone_lib="NCI_CGAP_Pan3"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.1 kb. Library constructed by Life Technologies."
      117 a 128 c 167 g 117 t

BASE COUNT 117 a 128 c 167 g 117 t
ORIGIN

alignment_scores:
  Quality: 137.00 Length: 116
  Ratio: 1.957 Gaps: 4
  Percent Similarity: 60.345 Percent Identity: 30.172

alignment_block:
  US-09-215-435-179 x BE047063/rev ..
  Align seg 1/1 to reverse of: BE047063 from: 1 to: 529
  9 LeuLeuPheLeuLeuPhePheLeuLeuLeu..... 20
  :::::::::::::::::::: |||||
  499 ATCTTGATCTCTCTGCTGCGCCCTCTTTGGCCTATGCAGCTCCTGG 450
  21 ThrArgGlySerLeuSerProThrLysTyrAsnLeuLeuLeuLysG 37
  ||||| :: |||||
  449 CCCCGGGGGATCATTTATCAAC.....CTGGAGACGGTG 415
  37 luserCysileArgAsnGlnAspCysGluThrGlyCysCysGlnArgAla 53
  ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  414 AGCTCTGCATGAATAGTGCCCGCCAGTGAAGAGCAATTGCTGCCAGCATTC 365
  54 ProAspAsnCysGluSerHisCysAlaGluLysGlySerGluGlySerLe 70
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  364 AGTCGCTGGCCCTGCGCCGCTGCACATCCATGCCGAGCAGACGCA 315
  70 uCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysProCysLeu 87
  ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  314 GTGCTGTCTAAGACGCTCTATGGGATTTACTACAGTGTCCCTGTGAGC 265
  87 rGAsnLeuThrCysileTyrSerLysAsnGluLysTrpLeu..... 100
  ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  264 GTGGCTGACCTGT.....GAGGAGACAGACCATCGTGGGCTCC 224
  101 .....SerIleAlaTyrGlyArgCysGlnLysileGlyArgGlnLys 114
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  223 ATCACCAACACCAACTTTGGCATCTGCCATGACCTGGAGCTCCAAG 176

seq_name: gb_est81:BE969660

seq_documentation_block: 528 bp mRNA EST 04-OCT-2000
LOCUS BE969660
DEFINITION 601679794F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3949990 5',
mRNA sequence.
ACCESSION BE969660
VERSION BE969660.1 GI:10582593
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 528)
  NTH-MGC http://mgi.nci.nih.gov/.
  TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL Unpublished (1999)
  CONTACT: Robert Strausberg, Ph.D.
  Email: cgapbs@mail.nih.gov
  Tissue Procurement: CLONTECH Laboratories, Inc.

```

```

cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC815 row: 1 column: 23
High quality sequence stop: 525.
Location/Qualifiers
  1..528
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:3949990"
  /clone_lib="NIH_MGC_78"
  /lab_host="DH10B (T1 phage-resistant)"
  /note="Organ: pancreas; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgctgcgc); Site_2: SfiI (ggccattagccc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCGGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
  117 a 168 c 133 g 110 t

BASE COUNT 117 a 168 c 133 g 110 t
ORIGIN

alignment_scores:
  Quality: 136.50 Length: 111
  Ratio: 1.923 Gaps: 3
  Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:
  US-09-215-435-179 x BE969660 ..
  Align seg 1/1 to: BE969660 from: 1 to: 528
  12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProTh 28
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  26 ATCTTGATCTCTCTGCTGCGCCCTCTCTGTGGCCCATGCAGCTCCTGG 75
  28 rLysTyrAsnLeu.....LeuGluLeuLysGluSerCysileArg 42
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  76 CCCCGGGGGATCATTTATCAACCTGGAGACGGTGAGCTGTGCATGAATA 125
  42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  126 GTGCCAGTGTAAAGACAAATTGCTCCAGCATTTCAAGTGGCGTGGCCCTG 175
  59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  176 GCCCGCTGCACATCCATGCCAGCAGCAGACAGCGAGTGTCTGTCAAGAC 225
  75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  226 GCTCTATGGGATTTACTACAAAGTGTCCCTGTGAGCGTGGCCTGACCTGT 274
  92 letyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  275 .....GAGGAGACAGACCATCGTGGGCTCCATCACCAACACCAAC 316
  104 TyrGlyArgCysGlnLysileGlyArgGlnLys 114
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
  317 TTTTGGCATCTGCCATGACCTGGAGCTCCAAG 349

seq_name: gb_est81:BE969913
seq_documentation_block: 533 bp mRNA EST 04-OCT-2000
LOCUS BE969913
DEFINITION 601680001F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950113 5',
mRNA sequence.

```

```

ACCESSION BE969913
VERSION BE969913.1 GI:10582846
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM816 row: b column: 02
High quality sequence stop: 487.
Location/Qualifiers
FEATURES
source
1. .533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950113"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggccattatggcc); Site_2: SfII (ggccattatggcc)
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

BASE COUNT 118 a 169 c 132 g 114 t

ORIGIN

alignment_scores:

Quality:	135.50	Length:	111
Ratio:	1.908	Gaps:	3
Percent Similarity:	63.964	Percent Identity:	28.829

alignment_block:

US-09-215-435-179 x BE969913 ..

Align seg 1/1 to: BE969913 from: 1 to: 533

```

12 LeuLeuPhePheLeuLeuLeuLeuThrArgGlySerLeuSerProTh 28
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
26 ATCTCTGATCCTCTGCTGTGCGCCCTCTCTGTGGCCTATGCAGCTCCTGG 75

28 rlystyrAsnLeu.....LeuGlulLeuLysGluSerCysIleArgA 42
:::|||||:|||||:|||||:|||||:|||||:|||||:
76 CCCCCTGGGGGATCATTTATCAACCTCGAGAACGGTGTGCTGCTGCAATA 125

42 snGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCysGlu 58
:::|||||:|||||:|||||:|||||:|||||:|||||:
126 GTGCCAGTGTAAAGACAAATGCTCCACGATTCAGTCCGCTGGCCCTG 175

59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
:::|||||:|||||:|||||:|||||:|||||:|||||:
176 GCCCGCTGCACATCCATGCCAGCCAGCAGAACGAGTGCTCTGTCAAGAC 225

75 lPhePheGlyGlnThrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
:::|||||:|||||:|||||:|||||:|||||:|||||:
226 GCTCTATGGGATTTACTACAGTGTCCCTGTGTGAGCGTGGCCTGACCTGT. 274
```

```

92 lEtyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
:::|||||:|||||:|||||:|||||:|||||:
275 .....GAGGAGACAAACCATCGTGGGCTCCATCACCAACCAAC 316

104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
:::|||||:|||||:|||||:|||||:|||||:
317 TTTGGCATCTGCCATGACGCTGGAGGCTGCAAG 349

seq_name: gb_est81:BE969784

seq_documentation_block:
LOCUS BE969784 543 bp mRNA EST 04-OCT-2000
DEFINITION 601679544F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3949859 5',
mRNA sequence.
ACCESSION BE969784
VERSION BE969784.1 GI:10582717
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM815 row: g column: 12
High quality sequence stop: 477.
Location/Qualifiers
FEATURES
source
1. .543
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949859"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggccgctggcc); Site_2: SfII (ggccattatggcc)
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

BASE COUNT 118 a 176 c 133 g 115 t

ORIGIN

alignment_scores:

Quality:	135.50	Length:	111
Ratio:	1.908	Gaps:	3
Percent Similarity:	63.964	Percent Identity:	28.829

alignment_block:

US-09-215-435-179 x BE969784 ..

Align seg 1/1 to: BE969784 from: 1 to: 543

```

12 LeuLeuPhePheLeuLeuLeuLeuThrArgGlySerLeuSerProTh 28
:::|||||:|||||:|||||:|||||:|||||:|||||:
26 ATCTCTGATCCTCTGCTGTGCGCCCTCTCTGTGGCCTATGCAGCTCCTGG 75

28 rlystyrAsnLeu.....LeuGlulLeuLysGluSerCysIleArgA 42
:::|||||:|||||:|||||:|||||:|||||:|||||:
126 GTGCCAGTGTAAAGACAAATGCTCCACGATTCAGTCCGCTGGCCCTG 175

59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
:::|||||:|||||:|||||:|||||:|||||:|||||:
176 GCCCGCTGCACATCCATGCCAGCCAGCAGAACGAGTGCTCTGTCAAGAC 225

75 lPhePheGlyGlnThrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
:::|||||:|||||:|||||:|||||:|||||:|||||:
226 GCTCTATGGGATTTACTACAGTGTCCCTGTGTGAGCGTGGCCTGACCTGT. 274
```

```

76  CCCCCGGGGGATCATTTATCAACTCGAGAACGGTGAGCTCTGCATGAATA 125
42  sNlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCysGlu 58
   :: :|||:::|||||:::
126 GTGCCAGCTCTAAGAGCAATTGCTGCCAGCATTTCAAAGTGCCTGGGCCCTG 175
59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
   :: :|||:::|||||:::|||||:::
176 GCCCGCTGCACATCCATCGCCGCGGAGAACGGAGTGTCTGTCTGCACAGAC 225
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
   :: :|||:::|||||:::|||||:::
226 GCTCATGGGATTTACTACAAGTGTCCCTGTGAGCGTGGCCTCACCTGT . 274
92 leTyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
   :: :|||:::|||||:::
275 .....GAGGGAGACAAGACCATCGTGGCGTCCATACCAACACCAAC 316
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
   :: :|||:::|||||:::
317 TTTGGCATCTGCCATGAGCGTGGACGCTCCAAG 349

seq_name: gb_est81:BE969920

seq_documentation_block:
LOCUS BE969920 577 bp mRNA
DEFINITION 601680009F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950115 5',
           mRNA sequence.
ACCESSION BE969920
VERSION BE969920.1 GI:10582853
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

1 (bases 1 to 577)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LiCM816 row: 6 column: 04
High quality sequence stop: 534.
Location/Qualifiers
1. 577
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950115"
/clone_l1b="NIH-MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgctcgccg); Site_2: SfII (ggcattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGAGCGGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
139 a 174 c 144 g 120 t
BASE COUNT
ORIGIN
alignment_scores:

```

); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGGCACATG-dT(30)BN-3', (where B = A, C, G and N = A, C, G, or T). Average insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 173 a 198 c 191 g 120 t
ORIGIN

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:
US-09-215-435-179 x BE969726 ..

Align seg 1/1 to: BE969726 from: 1 to: 682

12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProTh 28
26 ATCTGATCTCTCTGCTGTCGCCCTCTCTGTGGCTATGCAGCTCTGG 75
28 rlytyrAsnLeu.....LeuGlulLeuLysGluSerCysIleArgA 42
76 CCCCCTGGGGATCATATCAACCTGGAGACGGTGCAGCTCTGCATGAATA 125
42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
126 GTGCCAGTGAAGAGCAATTGCTGCCAGCAATCAAGTGCAGTGCAGCTG 175
59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
176 GCCCCGTGCACATCCATGACCGAGAGACAGCGAGTCTGTCAAGAC 225
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
226 GCTCATGGGATTACTACAAAGTGTCCCTGTGAGCGTGCCTGACCTGT. 274
92 leTyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
275GAGGAGACAGACCATCGTGGCTCCATCACCACCAACCAAC 316
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
317 TTGGCATCTGCCATGCAGCTGGAGCTCCCAAG 349

seq_name: gb_est81:BE969942

seq_documentation_block:
LOCUS BE969942 719 bp mRNA EST 04-OCT-2000
DEFINITION 601680032F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950455 5',
mRNA sequence.
ACCESSION BE969942
VERSION BE969942.1 GI:10582875
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 719)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: ccapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LiCM816 row: p column: 08
High quality sequence stop: 524.

FEATURES

source
1..719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950455"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (TI phage-resistant)"
/note="organ: pancreas; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggcgattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGGCACATG-dT(30)BN-3', (where B = A, C, G, or N = A, C, G, or T). Average insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 160 a 233 c 191 g 135 t
ORIGIN

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

alignment_block:

US-09-215-435-179 x BE969942 ..

Align seg 1/1 to: BE969942 from: 1 to: 719

12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProTh 28
26 ATCTGATCTCTCTGCTGTCGCCCTCTCTGTGGCTATGCAGCTCTGG 75
28 rlytyrAsnLeu.....LeuGlulLeuLysGluSerCysIleArgA 42
76 CCCCCTGGGGATCATATCAACCTGGAGACGGTGCAGCTCTGCATGAATA 125
42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
126 GTGCCAGTGAAGAGCAATTGCTGCCAGCAATCAAGTGCAGTGCAGCTG 175
59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
176 GCCCCGTGCACATCCATGACCGAGAGACAGCGAGTCTGTCAAGAC 225
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
226 GCTCATGGGATTACTACAAAGTGTCCCTGTGAGCGTGCCTGACCTGT. 274
92 leTyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
275GAGGAGACAGACCATCGTGGCTCCATCACCACCAACCAAC 316
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
317 TTGGCATCTGCCATGCAGCTGGAGCTCCCAAG 349

seq_name: gb_est81:BE970098

seq_documentation_block:

LOCUS BE970098 733 bp mRNA EST 04-OCT-2000
DEFINITION 601679841F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950098 5',
mRNA sequence.
ACCESSION BE970098
VERSION BE970098.1 GI:10583131
KEYWORDS EST.

```

SOURCE          human.
ORGANISM         Homo sapiens
REFERENCE        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE           1 (bases 1 to 733)
JOURNAL          NIH-MGC http://mgi.nci.nih.gov/.
COMMENT         Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: CLONTECH Laboratories, Inc.
                CDNA Library Preparation: CLONTECH Laboratories, Inc.
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: L1C8M16 row: a column: 11
                High quality sequence stop: 499.
FEATURES         Location/Qualifiers
                1..733
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:3950098"
                /clone_lib="NIH_MGC_78"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
                Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggccattagggc
                ); 5' and 3' adaptors were used in cloning as follows: 5'
                adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor
                sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dT(30)BN-3'
                (where B = A, C, or G and N = A, C, G, or T). Average
                insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
                contained inserts by PCR. This library was enriched for
                full-length clones and was constructed by Clontech
                Laboratories (Palo Alto, CA)."
                BASE COUNT      171 a 218 c 210 g 134 t
                ORIGIN
                alignment_scores:
                Quality: 135.50      Length: 111
                Ratio: 1.908          Gaps: 3
                Percent Similarity: 63.964      Percent Identity: 28.829
                alignment_block:
                US-09-215-435-179 x BE970098
                Align seg 1/1 to: BE970098 from: 1 to: 733
                12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProth 28
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                26 ATCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 75
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                28 rLysTyrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                76 CCCCCGGGGATCATTTATCAACCTGGAGAACGGTGAGCTGTGCATGAATA 125
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
                :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                126 GTGCCAGTGTAAAGCAATTTGCTGCCAGCATTCAGTGGCGCTGGCGCTG 175
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                59 SerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThrGlnVa 75
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                176 GCCCCCTGCATCATCCATCGCCAGGAGAACACCGAGTGTCTGTGCACAGAC 225
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                75 1PhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                226 GCTCATGGGATTTACTACAAGTGTCCCTGTGAGCGTGGCGTGCACCTGT. 274
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                92 letYrSerLysAsnGluLysTrpLeu.....SerIleAla 103
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
                275 .....GAGGGAGACAAGACCATCTGCTGGGCTCCATCACCACCAAC 316
                :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
317 TTGGGATCTGTCATGACGCTGGAGCGCTCCAA 349
seq_name: gb_est81:BE970031
seq_documentation_block:
LOCUS      BE970031      752 bp      mRNA      EST      04-OCT-2000
DEFINITION 601680144F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950266 5',
            mRNA sequence.
ACCESSION  BE970031
VERSION    BE970031.1 GI:10582964
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 752)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            CDNA Library Preparation: CLONTECH Laboratories, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1C8M16 row: h column: 11
            High quality sequence stop: 523.
FEATURES   Location/Qualifiers
            1..752
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3950266"
            /clone_lib="NIH_MGC_78"
            /lab_host="DH10B (T1 phage-resistant)"
            /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
            Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggccattagggc
            ); 5' and 3' adaptors were used in cloning as follows: 5'
            adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor
            sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dT(30)BN-3'
            (where B = A, C, or G and N = A, C, G, or T). Average
            insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
            contained inserts by PCR. This library was enriched for
            full-length clones and was constructed by Clontech
            Laboratories (Palo Alto, CA)."
            BASE COUNT      168 a 255 c 201 g 128 t
            ORIGIN
            alignment_scores:
            Quality: 135.50      Length: 111
            Ratio: 1.908          Gaps: 3
            Percent Similarity: 63.964      Percent Identity: 28.829
            alignment_block:
            US-09-215-435-179 x BE970031
            Align seg 1/1 to: BE970031 from: 1 to: 752
            12 LeuLeuPhePheLeuPheLeuLeuThrArgGlySerLeuSerProth 28
            :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
            26 ATCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 75
            :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
            28 rLysTyrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
            :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
            76 CCCCCGGGGATCATTTATCAACCTGGAGAACGGTGAGCTGTGCATGAATA 125
            :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
            42 snGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnCysGlu 58
            :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

126 GTGCCAGTGTAAAGACCAATTGCTCCAGCAATTCAAGTGCCTGGGCGTG 175
59 SerHisCysAlaGluLysSerGluCysGlnThrGlnVa 75
176 GCCCGCTGCATCCATCCAGCCAGCAGAACAGGAGTCTGTGTCAGAAC 225
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
226 GCTCTATGGGATTACTAAGAGTGCCTGTGAGCGTGGCGCTGACCTGT. 274
92 lTyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
275 .....GAGGAGACAAAGACCATCGTGGCTCCATCACCACCAACCA 316
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
317 TTTGGCATCTGCCATGAGCGTGGAGCTCCCAAG 349
seq_name: gb_est81:BE970121

seq_documentation_block:
LOCUS BE970121 764 bp mRNA EST 04-OCT-2000
DEFINITION 601679866F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950152 5',
mRNA sequence.
ACCESSION BE970121
VERSION BE970121.1 GI:10583154
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 764)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM816 row: c column: 17
High quality sequence stop: 465.

FEATURES
source
1..764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950152"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc)
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 175 a 241 c 211 g 137 t
ORIGIN

```

```

alignment_scores:
Quality: 135.50 Length: 111
Ratio: 1.908 Gaps: 3
Percent Similarity: 63.964 Percent Identity: 28.829

```

```

alignment_block:
US-09-215-435-179 x BE970121 ..
Align seg 1/1 to: BE970121 from: 1 to: 764
12 LeuLeuPhePhePheLeuLeuLeuThrArgGlySerLeuSerProTh 28
:::|||||:::|||||:::|||||:::|||||:::|||||
26 ATCTGTGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 75
28 rLysTyrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
:::|||||:::|||||:::|||||:::|||||:::|||||
76 CCCCGGGGGGATCATATCAACCTGGAGAACGGTGAAGTCTGCATGAATA 125
42 snGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCysGlu 58
:::|||||:::|||||:::|||||:::|||||:::|||||
126 GTGCCAGTGTAAAGACCAATTGCTCCAGCAATTCAAGTGCCTGGCGCTG 175
59 SerHisCysAlaGluLysGlySerGluCysGlnThrGlnVa 75
:::|||||:::|||||:::|||||:::|||||:::|||||
176 GCCCGCTGCATCCATCCAGCCAGCAGAACAGGAGTCTGTGTCAGAAC 225
75 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 92
:::|||||:::|||||:::|||||:::|||||:::|||||
226 GCTCTATGGGATTACTAAGAGTGCCTGTGAGCGTGGCGCTGACCTGT. 274
92 lTyrSerLysAsnGluLysTrpLeu.....SerIleAla 103
:::|||||:::|||||:::|||||:::|||||:::|||||
275 .....GAGGAGACAAAGACCATCGTGGCTCCATCACCACCAACCA 316
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
:::|||||:::|||||:::|||||:::|||||:::|||||
317 TTTGGCATCTGCCATGAGCGTGGAGCTCCCAAG 349
* seq_name: gb_est81:BE970243

seq_documentation_block:
LOCUS BE970243 772 bp mRNA EST 04-OCT-2000
DEFINITION 601679987F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950397 5',
mRNA sequence.
ACCESSION BE970243
VERSION BE970243.1 GI:10583088
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM816 row: m column: 22
High quality sequence stop: 518.

FEATURES
source
1..772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950397"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc)
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 175 a 241 c 211 g 137 t
ORIGIN

```

```

alignment_scores:      Quality: 135.50      Length:      111
                       Ratio:   1.908       Gaps:         3
Percent Similarity:    63.964     Percent Identity: 28.829

alignment_block:
US-09-215-435-179 x BE970243 ..

Align seg 1/1 to: BE970243 from: 1 to: 772

12 LeuLeuPhePhePheLeuPheLeuThrArgGlySerLeuSerProTh 28
   :|::||:::||::||::||::||::||::||::||::||::||::||
26 ATCTGATCTCCTGTGTTCGCCCTCTCTGTGGCCATTGCAGTCTCGTG 75
   :|::||::||::||::||::||::||::||::||::||::||::||
28 rLysTyrrAsnLeu.....LeuGluLeuLysGluSerCysIleArgA 42
   :|::||::||::||::||::||::||::||::||::||::||::||
76 CCCCCGGGGATCATATCAACTGGASACGGTGAGCTCTGCATGAATA 125
   :|::||::||::||::||::||::||::||::||::||::||::||
42 snGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnCysGlu 58
   :|::||::||::||::||::||::||::||::||::||::||::||
126 GTGCCACAGTTAAAGAGCAATTGTCGCCAGCATTCAAAGTCGCGCTG 175
   :|::||::||::||::||::||::||::||::||::||::||::||
59 SerHisCysAlaGluYsGlySerGluGlySerLeuCysGlnThrGlnVa 75
   :|::||::||::||::||::||::||::||::||::||::||::||
176 GCCCGCTGCACATCCATGCGCAGCGAGAACAACGACGAGTGCTCTGT 225
   :|::||::||::||::||::||::||::||::||::||::||::||
75 lPhePheGlyGlnTyrrArgAlaCysProCysLeuArgAsnLeuthrCy 92
   :|::||::||::||::||::||::||::||::||::||::||::||
226 GCTCTATGGGATTACTACAAGTGTCCCTGTGTGAGCGTGGCCTGACCT 274
   :|::||::||::||::||::||::||::||::||::||::||::||
92 leYrSerLysAsnGluLysnrPLeu.....SerileAla 103
   :|::||::||::||::||::||::||::||::||::||::||::||
275 .....GAGGGAGACANGACCATTCTGGGCTCCATCCCNACACCAC 316
   :|::||::||::||::||::||::||::||::||::||::||::||
104 TyrGlyArgCysGlnLysIleGlyArgGlnLys 114
   :|::||::||::||::||::||::||::||::||::||::||::||
117 TTTCGGCATCTCCCATGAGCTGTGAGCGCTCCAAG 349

```


OM of: US-09-215-435-179_COPY_1_23 to: GenEmbl.* out_format : pfs

Date: Nov 15, 2001 12:45 AM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=framet-p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -QEXT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435_CGNL_1_11006 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-215-435-179_COPY_1_23

Query length: 23

Database: GenEmbl.*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 7698.940000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
gb_pr6:AL157823	+	119.00	221.87	0.0004	127384 ! AL157823 Human DNA sequence
gb_btg22:AL512371	+	119.00	220.11	0.0005	159482 ! AL512371 Homo sapiens chrom
gb_btg23:AL590400	+	119.00	219.66	0.0005	168884 ! AL590400 Homo sapiens chrom
gb_pr2:AC007056	+	70.00	111.75	548.17	182481 ! AC007056 Homo sapiens clone
gb_btg10:AC026232	+	68.50	108.04	881.75	192561 ! AC026232 Mus musculus chrom
gb_btg14:AC068663	-	68.50	107.26	975.41	212308 ! AC068663 Mus musculus chrom
gb_btg24:AP002376	-	68.00	108.36	846.35	160759 ! AP002376 Homo sapiens chrom
gb_btg18:AC087103	+	68.00	107.85	903.97	171648 ! AC087103 Papio hamadryas cld
gb_btg20:AL158168	+	68.00	107.29	970.60	184233 ! AL158168 Homo sapiens chrom
gb_pr1:AC004644	+	67.00	117.20	272.34	39338 ! AC004644 Homo sapiens chrom
gb_btg25:HSDJ821D9	+	67.00	107.14	989.34	141991 ! AL121960 Homo sapiens chrom
gb_pr8:HS240418	+	67.00	106.50	1.1e+03	154066 ! AL021808 Human DNA sequence
gb_btg3:AC012220	+	67.00	106.26	1.1e+03	158947 ! AC012220 Homo sapiens chrom
gb_btg2:AC009762	-	67.00	105.29	1.3e+03	179956 ! AC009762 Homo sapiens clone
gb_pr9:HSSDH855	+	66.00	147.18	5.83	649 ! U17883 Human succinate dehydro
gb_btg24:AP003180	+	66.00	104.33	1.4e+03	153814 ! AP003180 Homo sapiens chrom
gb_btg19:AL138896	+	66.00	104.20	1.4e+03	156375 ! AL138896 Homo sapiens chrom
gb_btg24:AP003401	+	66.00	104.07	1.5e+03	158977 ! AP003401 Homo sapiens chrom
gb_btg7:AC021317	+	66.00	103.79	1.5e+03	164847 ! AC021317 Homo sapiens clone
gb_btg5:AC020746	-	66.00	103.70	1.5e+03	166566 ! AC020746 Homo sapiens chrom
gb_btg5:AC015958	+	66.00	103.69	1.5e+03	166960 ! AC015958 Homo sapiens chrom
gb_btg10:AC025335	+	66.00	103.58	1.6e+03	169291 ! AC025335 Homo sapiens chrom
gb_btg19:AL139404	-	66.00	103.50	1.6e+03	170889 ! AL139404 Homo sapiens chrom
gb_btg24:AP003356	+	66.00	103.49	1.6e+03	171154 ! AP003356 Homo sapiens chrom
gb_btg8:AC022245	+	66.00	103.32	1.6e+03	174937 ! AC022245 Homo sapiens chrom
gb_btg14:AC068410	+	66.00	103.20	1.6e+03	177534 ! AC068410 Homo sapiens chrom
gb_btg18:AC090016	+	66.00	103.17	1.6e+03	178360 ! AC090016 Homo sapiens chrom
gb_btg21:ALJ56142	+	66.00	103.08	1.7e+03	180282 ! AL356142 Homo sapiens chrom
gb_btg22:ALJ39061	+	66.00	103.02	1.7e+03	181674 ! AL330961 Homo sapiens chrom
gb_btg10:AC026021	+	66.00	102.83	1.7e+03	186248 ! AC026021 Homo sapiens chrom
gb_btg23:AP001095	+	66.00	102.72	1.7e+03	188818 ! AP001095 Homo sapiens chrom
gb_btg14:AC067794	-	66.00	102.62	1.8e+03	191379 ! AC067794 Homo sapiens chrom
gb_btg24:AP002851	+	66.00	102.57	1.8e+03	192465 ! AP002851 Homo sapiens chrom
gb_btg18:AC087138	+	66.00	102.33	1.8e+03	198578 ! AC087138 Mus musculus clone
gb_btg24:AP002756	+	66.00	102.06	1.9e+03	205430 ! AP002756 Homo sapiens chrom
gb_btg13:AC063965	+	66.00	102.01	1.9e+03	206848 ! AC063965 Homo sapiens chrom
gb_btg23:AP001011	+	66.00	101.96	1.9e+03	207949 ! AP001011 Homo sapiens chrom
gb_btg24:AP002757	-	66.00	101.84	2.0e+03	211325 ! AP002757 Homo sapiens chrom
gb_pr9:HSDJ37C10	+	66.00	101.65	2.0e+03	216497 ! AL049569 Human DNA sequence
gb_pr4:AF067844	+	66.00	101.58	2.0e+03	218336 ! AF067844 Homo sapiens chrom
gb_btg6:AC018700	-	66.00	101.49	2.0e+03	228088 ! AC018700 Homo sapiens chrom

```
gb_btg3:DDIRAP1A - 65.00 136.22 23.77 1987 ! L09750 Dictyostelium discoi
gb_btg4:AC013976 - 65.00 124.07 112.89 9365 ! AC013976 Drosophila melanog
gb_pr8:HS206D15 + 65.00 105.39 1.2e+03 101574 ! AL021068 Homo sapiens DNA
gb_btg19:AL139250 + 65.00 104.98 1.3e+03 106962 ! AL139250 Homo sapiens chr
```

seq_name: gb_pr6:AL157823

```
seq_documentation_block:
LOCUS AL157823 127384 bp DNA PRI 10-FEB-2001
DEFINITION Human DNA sequence from clone RP3-51008 on chromosome 6, complete
sequence.
```

ACCESSION AL157823

VERSION AL157823.9 GI:12750800

KEYWORDS HTG.

SOURCE human.

ORGANISM

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 127384)
```

REFERENCE

AUTHORS Williams,S.

TITLE Direct Submission

JOURNAL Submitted (10-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerquest@sanger.ac.uk

On Feb 12, 2001 this sequence version replaced gi:12709869.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em; EMBL; Sw;

SWISSPROT; Tr; TREMBL; Wp; WORMPEP; information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6

RP3-51008 is from the library RP3-3 constructed by the group of

Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone

RP3-51008 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone RP3-51008 is at 1 in this sequence. The

true left end of clone RP3-422H11 is at 127285 in this sequence.

FEATURES

Location/Qualifiers

source

1..127384

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RP3-51008"

/clone_lib="RPC1-3"

6404

/note="Weak data in alu."

complement(95875..96037)

/note="match: STS: Em:G13580"

95894..96129

/note="match: STS: Em:G14863"

BASE COUNT 33605 a 30130 c 31408 g 32241 t

ORIGIN

```

/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-238F24"
/clone_lib="RPC1-11.1"
1..3066
/note="assembly_fragment:02384
fragment_chain:1"
3167..155002
/note="assembly_fragment:02553
fragment_chain:1"
155103..159482
/note="assembly_fragment:02886"
BASE COUNT 42530 a 37388 c 38532 g 40831 t 201 others
ORIGIN

alignment_scores:
  Quality: 119.00      Length: 23
  Ratio: 5.174         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x AL512371 ...
Align seg 1/1 to: AL512371 from: 1 to: 159482
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLe 17
95543 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTTCTTCTTCTTCT 95592
|||||
17 upheLeuLeuThrArgGly 23
|||||
95593 CTTCTCTCTCACCAGGGGC 95611

seq_name: gb_htg23:AL590400

seq_documentation_block:
LOCUS AL590400 168884 bp DNA HTG 08-APR-2001
DEFINITION Homo sapiens chromosome 6 clone RP11-569112, *** SEQUENCING IN
PROGRESS ***, 14 unordered pieces.
ACCESSION AL590400.2 GI:13568356
VERSION AL590400.2 GI:13568356
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 168884)
Sims,S.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On Apr 9, 2001 this sequence version replaced gi:13559131.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA569112
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:13559131.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA569112
----- Summary Statistics
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 154970 bases at least Q40
Consensus quality: 156246 bases at least Q30
Consensus quality: 156738 bases at least Q20
Insert size: 159282; sum-of-contigs
Insert size: 165487; 7.5% error; agarose-fp
Quality coverage: 6.35x in Q20 bases; sum-of-contigs Quality
coverage: 6.11x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3066: contig of 3066 bp in length
* 3067 3166: gap of 100 bp
* 3167 155002: contig of 151836 bp in length
* 155003 155102: gap of 100 bp
* 155103 159482: contig of 4380 bp in length.
*
* Location/Qualifiers
1..159482
FEATURES
source

```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1      2920: contig of 2920 bp in length
* 2921 3020: gap of 100 bp
* 3021 32168: contig of 29148 bp in length
* 32169 32268: gap of 100 bp
* 32269 36080: contig of 3812 bp in length
* 36081 36180: gap of 100 bp
* 36181 45704: contig of 9524 bp in length
* 45705 45804: gap of 100 bp
* 45805 59680: contig of 13876 bp in length
* 59681 59780: gap of 100 bp
* 59781 61890: contig of 2110 bp in length
* 61891 61990: gap of 100 bp
* 61991 72185: contig of 10195 bp in length
* 72186 72285: gap of 100 bp
* 72286 76243: contig of 3958 bp in length
* 76244 76343: gap of 100 bp
* 76344 81840: contig of 5497 bp in length
* 81841 81940: gap of 100 bp
* 81941 86241: contig of 4301 bp in length
* 86242 86341: gap of 100 bp
* 86342 129124: contig of 42783 bp in length
* 129125 129224: gap of 100 bp
* 129225 131985: contig of 2761 bp in length
* 131986 132085: gap of 100 bp
* 132086 151439: contig of 19354 bp in length
* 151440 151539: gap of 100 bp
* 151540 168884: contig of 17345 bp in length.

```

FEATURES

```

source
1. .168884
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="6"
  /clone="RP11-569I12"
  /clone_lib="RPC1-11.2"
1. .2920
  /note="assembly_fragment:00514"
  fragment_chain:1
  clone_end:SP6
  vector_side:left
3021..32168
  /note="assembly_fragment:01440"
  fragment_chain:1
32269..36080
  /note="assembly_fragment:00777"
  fragment_chain:1
36181..45704
  /note="assembly_fragment:00971"
  fragment_chain:1
45805..59680
  /note="assembly_fragment:00349"
  fragment_chain:1
59781..61890
  /note="assembly_fragment:02150"
  fragment_chain:2
61991..72185
  /note="assembly_fragment:02579"
  fragment_chain:2
72286..76243
  /note="assembly_fragment:00659"
  fragment_chain:2
76344..81840
  /note="assembly_fragment:00425"
  fragment_chain:2
81941..86241
  /note="assembly_fragment:02764"

```

```

misc_feature
86342..129124
  /note="assembly_fragment:00846"
  fragment_chain:2
129225..131985
  /note="assembly_fragment:01963"
  fragment_chain:3
132086..151439
  /note="assembly_fragment:02120"
  fragment_chain:3
151540..168884
  /note="assembly_fragment:01758"
  fragment_chain:3
  clone_end:T7
  vector_side:right"
BASE COUNT 46011 a 38577 c 39446 g 43544 t 1306 others
ORIGIN

alignment_scores
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block
US-09-215-435-179_COPY_1_23 x AL590400 ..
Align seg 1/1 to: AL590400 from: 1 to: 168884

1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLeu 17
|||||
109039 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTTCTTCTTCTTCT 109088

17 uPheLeuLeuThrArgGly 23
|||||
109089 CTTCTCTCTCACCGGGC 109107

seq_name: gb_pr2:AC007056

seq_documentation_block:
LOCUS AC007056 182481 bp DNA PRI 01-JUL-1999
DEFINITION Homo sapiens clone from human chromosome 14q31 region containing
gene for neurexin III, partial CDS, complete sequence.
ACCESSION AC007056
VERSION AC007056.4 GI:5306227
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 182481)
Young,J., Rowen,L., Madan,A., Qin,S., Abbasi,N., Dors,M., Dahl,T.,
Dickhoff,R., Hall,J., James,R., Loretz,C., Lasky,S., Madan,A.,
Prescott,S., Ratcliffe,A., Shaffer,T. and Hood,L.
Sequencing of human chromosome 14 gene for neurexin III
Unpublished
2 (bases 1 to 182481)
Young,J.M., Qin,S., Madan,A., Abbasi,N., Madan,A., Dors,M.,
James,R., Loretz,C., Lasky,S., Rowen,L. and Hood,L.
Direct Submission
Submitted (11-MAR-1999) Multimegabase Sequencing Group, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
3 (bases 1 to 182481)
Young,J., Rowen,L., Madan,A., Qin,S., Abbasi,N., Dors,M., Dahl,T.,
Dickhoff,R., Hall,J., James,R., Loretz,C., Lasky,S., Madan,A.,
Prescott,S., Ratcliffe,A., Shaffer,T. and Hood,L.
Direct Submission
Submitted (01-JUL-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Jul 1, 1999 this sequence version replaced gi:4966341.
This sequence overlaps BAC 332E19, found in GenBank Accession
Number AF099810.
FEATURES
Location/Qualifiers
1. .182481

```

Thu Nov 15 10:52:02 2001

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q31"
/clone="BAC 31E9"
/clone_lib="RPCI-11"
1. 23236
misc_feature
note="Overlap span with BAC 332E19, found in AF099810, at
positions 170589-193824."
1702
variation
/replace="g"
/replace="g"
3369
variation
/replace="c"
3867..3869
variation
/replace="t"
4378
variation
/replace="gccag"
4858
variation
/replace="a"
5277
variation
/replace="t"
5393
variation
/replace="g"
6561
variation
/replace="g"
8235
variation
/replace="c"
8591
variation
/replace="a"
10516
variation
/replace="t"
11378
variation
/replace="g"
13958
variation
/replace="g"
13994
variation
/replace="g"
14454
variation
/replace="c"
14518
variation
/replace="a"
14619
variation
/replace="t"
14679
variation
/replace="g"
14753
variation
/replace="g"
15144
variation
/replace="g"
15809
variation
/replace="c"
16109
variation
/replace="c"
/replace="g"
16639
variation
/replace="g"
16725
variation
/replace="c"
17208
variation
/replace="g"
17311
variation
/replace="c"
17579
variation
/replace="g"
17731
variation
/replace="c"
18114
variation
/replace="c"
18491
variation
/replace="g"
18728
variation
/replace="c"
19388
variation
/replace="a"
19728
variation
/replace="g"
20250
variation
/replace="g"
21522
variation
/replace="a"
21838..21839
variation
/replace="a"
21902
variation
/replace="g"
21978
variation
/replace="c"
22190
variation
/replace="a"
22265
variation
/replace="a"
22355
variation
/replace="g"
22525..22526
variation
/replace="aa; 332E19: a"
22924
variation
/replace="t; 332E19: a"
59442..69550
unsure
/replace="a"
69650..69762
unsure
/replace="low quality data"
<75358..>75426
CDS
/replace="Matches the beginning of KIAA0743, found in
AB018286. The start position of the exon is unclear."
/codon_start=3

```


Thu Nov 15 10:52:02 2001

```

WORKING DRAFT SEQUENCE, 25 unordered pieces.
AC068663
AC068663.2 GI:8134859
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 212908)
Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,Q.L., Maduro,V.B.,
Mastrian,S.D., McCloskey,J.C., Morse,E., Ojoda,M.A., Pearson,R.,
Stantipop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
Wetherby,K.D. and Green,E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 212908)
Green,E.D.
Direct Submission
Submitted (06-MAY-2000) NIH Intramural Sequencing Center, 8717
Groveomt Circle, Gaithersburg, MD 20877 USA
On Jun 1, 2000 this sequence version replaced gi:7712112.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_mouse@hgri.nih.gov
----- Project Information
Center project name: x1
Center clone name: 280D18
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 19975 bases at least Q40
Consensus quality: 20277 bases at least Q30
Consensus quality: 204027 bases at least Q20
Insert size: 222000; agarose-fp
Insert size: 216000; pulse-field-gel
Insert size: 210508; sum-of-contigs
Quality coverage: 4.37x in Q20 bases; agarose-fp
Quality coverage: 4.49x in Q20 bases; pulse-field-gel
Quality coverage: 4.61x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2371: contig of 2371 bp in length
* 2372 2471: gap of unknown length
* 2472 4622: contig of 2151 bp in length
* 4623 4722: gap of unknown length
* 4723 7917: contig of 3194 bp in length
* 7917 8016: gap of unknown length
* 8017 10417: contig of 2401 bp in length
* 10418 10517: gap of unknown length
* 10518 14260: contig of 3743 bp in length
* 14261 14361: gap of unknown length
* 14361 18020: contig of 3660 bp in length
* 18021 18121: gap of unknown length
* 18121 22446: contig of 4325 bp in length
* 22446 22546: gap of unknown length
* 22546 26537: contig of 3992 bp in length
* 26538 26638: gap of unknown length
* 26638 30194: contig of 3557 bp in length
* 30195 30295: gap of unknown length
* 30295 34805: contig of 4511 bp in length
*
* 34806 34905: gap of unknown length
* 34906 39674: contig of 4769 bp in length
* 39675 44570: contig of 4796 bp in length
* 44571 49670: gap of unknown length
* 49671 50069: contig of 5299 bp in length
* 50070 55626: contig of 5556 bp in length
* 55627 57225: gap of unknown length
* 57226 61503: contig of 5777 bp in length
* 61504 70931: gap of unknown length
* 70932 79217: contig of 8186 bp in length
* 79218 88380: contig of 9063 bp in length
* 88381 102624: contig of 14144 bp in length
* 102625 112658: contig of 9934 bp in length
* 112659 130269: gap of unknown length
* 130270 149318: contig of 17411 bp in length
* 149319 170176: gap of unknown length
* 170177 189732: contig of 19049 bp in length
* 189733 212908: contig of 20658 bp in length
* 212909 212908: contig of 19556 bp in length
* 212909 212908: contig of 23076 bp in length.
*
* 212908
* /organism="Mus musculus"
* /strain="C57BL6/J"
* /db_xref="taxon:10090"
* /chromosome="5"
* /clone_lib="RPCI mouse BAC library 23"
* /clone="RP23-280D18"
* /note="assembly_fragment"
* 1. .2371
* /note="assembly_fragment"
* 2472. .4622
* /note="assembly_fragment"
* 4723. .7916
* /note="assembly_fragment"
* 8017. .10417
* /note="assembly_fragment"
* clone_end:SP6
* vector_side:right"
* 10518. .14260
* /note="assembly_fragment"
* 14361. .18020
* /note="assembly_fragment"
* 18121. .22445
* /note="assembly_fragment"
* 22546. .26537
* /note="assembly_fragment"
* 26638. .30194
* /note="assembly_fragment"
* 30295. .34805
* /note="assembly_fragment"
* 34906. .39674
* /note="assembly_fragment"
* 39775. .44570
* /note="assembly_fragment"
* 44671. .49969
* /note="assembly_fragment"
* 50070. .55625
* /note="assembly_fragment"
* 55726. .61502
* /note="assembly_fragment"
* 61603. .70931
* /note="assembly_fragment"
* 71032. .79217
*
* 212908
* /organism="Mus musculus"
* /strain="C57BL6/J"
* /db_xref="taxon:10090"
* /chromosome="5"
* /clone_lib="RPCI mouse BAC library 23"
* /clone="RP23-280D18"
* /note="assembly_fragment"
* 1. .2371
* /note="assembly_fragment"
* 2472. .4622
* /note="assembly_fragment"
* 4723. .7916
* /note="assembly_fragment"
* 8017. .10417
* /note="assembly_fragment"
* clone_end:SP6
* vector_side:right"
* 10518. .14260
* /note="assembly_fragment"
* 14361. .18020
* /note="assembly_fragment"
* 18121. .22445
* /note="assembly_fragment"
* 22546. .26537
* /note="assembly_fragment"
* 26638. .30194
* /note="assembly_fragment"
* 30295. .34805
* /note="assembly_fragment"
* 34906. .39674
* /note="assembly_fragment"
* 39775. .44570
* /note="assembly_fragment"
* 44671. .49969
* /note="assembly_fragment"
* 50070. .55625
* /note="assembly_fragment"
* 55726. .61502
* /note="assembly_fragment"
* 61603. .70931
* /note="assembly_fragment"
* 71032. .79217

```


seq name: qb_htq20:AL158168

```

seq_documentation_block:
LOCUS       AC087103       171648 bp       DNA              HTG          07-DEC-2000
DEFINITION  Papio hamadryas clone RP41-353B7, WORKING DRAFT SEQUENCE, 6
            unordered pieces.
ACCESSION   AC087103
VERSION     AC087103.1  GI:11596994
KEYWORDS    HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE      baboon
ORGANISM    Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
            Cercopithecinae; Papio.
REFERENCE   1 (bases 1 to 171648)
AUTHORS    Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            NISC Comparative Sequencing Initiative
            Unpublished
            2 (bases 1 to 171648)
            Green, E.D.
            Direct Submission
            Submitted (07-DEC-2000) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
            ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.mouse@nih.nih.gov
            ----- Project Information
            Center project name: asb
            Center clone name: 353B07
            ----- Summary Statistics
            Sequencing vector: plasmid; n/a; 100% of reads
            Sequencing strategy: shotgun; n/a; 100% of reads
            Sequencing date: 07-DEC-2000
            Sequencing center: NIH Intramural Sequencing Center
            Sequencing project: 353B07
            Sequencing project name: asb
            Sequencing project description: Papio hamadryas clone RP41-353B7, WORKING DRAFT SEQUENCE, 6
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q.,
            Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masiello, C.,
            Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
            Shevchenko, Y., Snyder, B., Stantirpop, S., Thomas, J.W., Thomas, P.J.,
            Tionsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
            Wetherby, K.D., Zhang, L.-H. and Green, E.D.
            Sequencing project title: NISC Comparative Sequencing Initiative
            Sequencing project accession: AC087103
            Sequencing project version: AC087103.1
            Sequencing project keywords: HTG: HTGS_PHASE1; HTGS_DRAFT.
            Sequencing project source: baboon
            Sequencing project organism: Papio hamadryas
            Sequencing project reference: 1 (bases 1 to 171648)
            Sequencing project authors: Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
            Bouffard, G.C., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,
            Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-
```


Thu Nov 15 10:52:02 2001

```

repeat_region complement(4172..4490)
/rpt_family="Alu"
repeat_region 4865..5171
/rpt_family="Alu"
repeat_region 5115..5147
/note="(A)33"
/rpt_type=tandem
/rpt_unit=A
repeat_region 5175..5317
/rpt_family="Alu"
repeat_region complement(5713..6019)
/rpt_family="Alu"
repeat_region 6123..6142
/note="(T)20"
/rpt_type=tandem
/rpt_unit=T
repeat_region complement(6123..6418)
/rpt_family="Alu"
repeat_region 6567..6850
/rpt_family="Alu"
repeat_region complement(7198..7496)
/rpt_family="Alu"
repeat_region complement(7387..7804)
/rpt_family="Alu"
misc_feature complement(8691..8782)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 8863..8967
/rpt_family="Alu"
repeat_region 9063..9153
/rpt_family="MIR"
repeat_region 9403..9757
/rpt_family="Alu"
repeat_region 9733..9757
/note="(A)25"
/rpt_type=tandem
/rpt_unit=A
repeat_region complement(9825..10147)
/rpt_family="Alu"
repeat_region complement(10413..10730)
/rpt_family="Alu"
repeat_region 11034..11292
/rpt_family="Alu"
repeat_region complement(11360..11685)
/rpt_family="Alu"
repeat_region 11360..11381
/note="(T)22"
/rpt_type=tandem
/rpt_unit=T
misc_feature 11701..11949
/note="100% identity EST ab01b06.sl"
/db_xref=dbEST:AA491442
repeat_region complement(12256..12850)
/rpt_family="Alu"
repeat_region complement(13025..13392)
/rpt_family="Alu"
repeat_region 13025..13051
/note="(T)27"
/rpt_type=tandem
/rpt_unit=T
repeat_region 13509..13648
/rpt_family="L1"
repeat_region 13656..13682
/note="(T)27"
/rpt_type=tandem
/rpt_unit=T
repeat_region complement(13762..13864)
/rpt_family="MIR"
misc_feature 14267..16252
/note="GRAIL 2 excellent exon, frame 1"
prim_transcript complement(15185..15716)
/standard_name=fmc16
/note="96% identity dbEST:AA631949"
repeat_region complement(16226..16423)
/rpt_family="MLT1"
complement(16476..16818)
/rpt_family="Alu"
17364..17450
/note="GRAIL 2 excellent exon, frame 0"
complement(17809..18097)
/rpt_family="Alu"
18194..18517
/rpt_family="Alu"
complement(18699..19020)
/rpt_family="Alu"
19197..19533
/rpt_family="Alu"
19724..19786
/rpt_family="Alu"
20071..20377
/rpt_family="Alu"
20544..20569
/note="(TG)13"
/rpt_type=tandem
/rpt_unit=TG
complement(20796..21116)
/rpt_family="Alu"
22158..22241
/note="GRAIL 2 excellent exon, frame 0"
22312..22657
/rpt_family="Alu"
complement(22798..23106)
/rpt_family="Alu"
22798..22833
/note="(T)36"
/rpt_type=tandem
/rpt_unit=T
23181..23228
/note="(ATTT)12"
/rpt_type=tandem
/rpt_unit=ATTT
23236..23571
/rpt_family="Alu"
complement(23575..23897)
/rpt_family="L1MB7"
complement(23893..24208)
/rpt_family="Alu"
complement(24207..24547)
/rpt_family="L1"
24827..24910
/note="GRAIL 2 excellent exon, frame 2"
25089..25441
/rpt_family="Alu"
complement(25787..25965)
/rpt_family="MER67C"
complement(26193..26492)
/rpt_family="Alu"
26873..26956
/note="GRAIL 2 excellent exon, frame 2"
27128..27171
/note="(ATGG)11"
/rpt_type=tandem
/rpt_unit=ATGG
27534..27561
/note="(ATGG)7"
/rpt_type=tandem
/rpt_unit=ATGG
27882..27950
/note="75% identity"
/rpt_type=tandem
/rpt_unit=TGA
28667..28750
/note="GRAIL 2 excellent exon, frame 2"
29384..29467
/note="GRAIL 2 excellent exon, frame 2"
29701..29799
/note="96% identity EST nj24c09.sl"

```


seq_documentation_block:
LOCUS HS24018 154066 bp DNA PRI 15-JUN-2000
DEFINITION Human DNA sequence from clone XHbac-24018 on chromosome 6p21.31-22.2 Contains zinc finger protein pseudogene, VNO-type olfactory receptor pseudogene, the gene for PRSS16 (protease, serine, 16 (thymus)), the gene for nuclear envelope pore membrane protein, ESTs, STSS, and GSSs, complete sequence.
ACCESSION AL021808
VERSION AL021808.1 GI:3395513
KEYWORDS Htg; nuclear envelope pore membrane protein; olfactory receptor; protease; PRSS16; zinc finger.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 154066)
Phillips,S.
Direct Submission
Submitted (23-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 5, 1998 this sequence version replaced gi:3355607.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C/elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
XHbac-24018 is from the Research Genetics total human BAC library that was screened by D. Ruddy and A. Gnirke from Mercator Genetics Inc. as published in Lauer et al (1997) Gen Res. 7: 457-470. This sequence is the entire insert of clone XHbac-24018.
FEATURES
source
1..154066
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="6"
/map="p21.31-22.2"
/clone_lib="SCB"
/clone="XHbac-24018"
1..1043
/notes="MCR52A repeat: matches 217. .1387 of consensus"
7429..7472
/note="22 copies 2 mer tt 77% conserved"
complement(7815..8187)
/notes="match: GSS: Em:AQ110321"
8219..8655
/notes="match: GSS: Em:AQ338654"
8618..11014
/gene="b24018.1"
8618..>11014
/gene="b24018.1"
/notes="b24018.1 (zinc finger)
match: cDNAs: Em:U66561 Em:LJ11672 Em:AF033199 Em:AF024693 Em:X59244 Em:AF020591 Em:X74855 Em:M27878 Em:D31763 Em:AF011573 Em:U75454 Em:M55422 Em:AF116030 Em:D89928 Em:M29580 Em:X17617 Em:U35376 Em:D70831 Em:X78925 Em:AB018341 Em:AF086368 Em:Y00850 Em:AB014528 Em:AB007872

Em:X65233 Em:M21977 Em:U78722 Em:L77247 Em:U09413
Em:X78933 Em:U38904 Em:L32163 Em:X78926 Em:U71599
match: ESTs: Em:T40438 Em:AI650923 Em:T29508 Em:T39154
Em:AA03845 Em:AA716653 Em:AA635575 Em:AA635567
Em:AI453500 Em:T90597 Em:D31299
match: proteins: Sw:P08045 Tr:O60792 Sw:Q99676 Sw:O75820
Tr:O75802 Sw:Q16587 Sw:P51523 Tr:Q14587 Tr:O62512
Sw:P51522 Tr:O43345 Sw:Q03923 Tr:O14585 Tr:O60765
Sw:Q07230 Sw:Q61751 Sw:P52742 Sw:P18753 Tr:O62886
Sw:P13682 Sw:P51786 Sw:P52738 Sw:P18739 Sw:Q62981
Tr:Q61116 Tr:Q14845 Tr:O75123 Tr:P70591 Sw:P51814
Sw:Q16600 Tr:Q15776 Sw:P10076 Tr:Q14586 Tr:O13398
Sw:Q02975 Wp:CE16599 Sw:O43830 Sw:Q62513 Tr:O43339
Sw:P21506 Tr:O43167 Sw:P15822 Sw:P51508 Sw:P18751
Sw:Q14929"
/codon_start=1
/evidence=not_experimental
complement(8897..9335)
/note="match: GSS: Em:AQ423094"
complement(8977..9310)
/note="match: GSS: Em:AQ461858"
complement(9925..9966)
/note="match: GSS: Em:AQ793807"
10443..10483
/gene="b24018.1"
/note="match: GSS: Em:AQ376888"
complement(10536..10666)
/note="match: GSS: Em:U77516"
complement(10689..11016)
/note="match: STS: Em:G11374"
complement(10854..11016)
/note="match: STS: Em:G14684"
complement(11017..11139)
/note="match: GSS: Em:U94452"
12877..13020
/note="L1MA4A repeat: matches 6119. .6280 of consensus"
13021..13761
/note="Tigger3(Golem) repeat: matches 2277. .3027 of consensus"
13753..13859
/note="Tigger3(Golem) repeat: matches 2. .108 of consensus"
13860..15869
/note="L1MA4A repeat: matches 4267. .6122 of consensus"
15870..16491
/note="LTR14B repeat: matches 5. .608 of consensus"
16492..19102
/note="L1MA4A repeat: matches 1857. .4267 of consensus"
19402..21319
/note="L1MA4A repeat: matches -675. .1857 of consensus"
22197..22580
/note="match: GSS: Em:AQ799446"
22697..22798
/note="51 copies 2 mer at 78% conserved"
complement(23267..23728)
/note="match: GSS: Em:AQ554556"
complement(23715..24097)
/note="match: GSS: Em:AQ266585"
25656..26249
/note="L1ME3 repeat: matches 4489. .5132 of consensus"
26250..26531
/note="AluSc repeat: matches 1. .282 of consensus"
26532..26785
/note="L1ME3 repeat: matches 5132. .5374 of consensus"
27386..27650
/note="MLTIG repeat: matches 10. .269 of consensus"
27671..27793
/note="L1ME3 repeat: matches 5403. .5511 of consensus"
28078..28455
/note="L1ME3 repeat: matches 5511. .5950 of consensus"
29086..29507
/note="match: GSS: Em:AQ033258"

us-09-215-435-179_copy_1_23.rge

Thu Nov 15 10:52:02 2001

...

OM of: US-09-215-435-179_COPY_1_23 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:28 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framed_p2n.model -DRV=xlp
-O=/cgml_1/USPTO_spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09215435_@CGML_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179_COPY_1_23

Query length: 23

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cgml_9/gcgdata/geneseq/geneseq/NA199.DAT:AAV51465 +	119.00	269.49	4.9e-07	421	
/cgml_9/gcgdata/geneseq/geneseq/NA199.DAT:AAV51465 +	119.00	267.31	6.4e-07	542	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	119.00	265.36	8.3e-07	680	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	62.50	140.73	7.22	474	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	62.00	136.31	12.73	695	
/cgml_9/gcgdata/geneseq/geneseq/NA194.DAT:AAV51465 +	62.00	131.16	24.65	1264	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	61.00	133.92	17.28	705	
/cgml_9/gcgdata/geneseq/geneseq/NA195.DAT:AAV51465 +	60.50	125.52	50.80	1641	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	60.00	141.31	6.70	230	
/cgml_9/gcgdata/geneseq/geneseq/NA198.DAT:AAV51465 +	59.50	135.43	14.24	399	
/cgml_9/gcgdata/geneseq/geneseq/NA197.DAT:AAV51465 +	59.00	122.67	73.16	1540	
/cgml_9/gcgdata/geneseq/geneseq/NA197.DAT:AAV51465 +	59.00	117.47	142.55	2817	
/cgml_9/gcgdata/geneseq/geneseq/NA197.DAT:AAV51465 +	59.00	112.06	285.51	5883	
/cgml_9/gcgdata/geneseq/geneseq/NA199.DAT:AAV51465 +	58.00	134.41	16.23	303	
/cgml_9/gcgdata/geneseq/geneseq/NA196.DAT:AAV51465 +	58.00	113.59	234.42	3399	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	58.00	90.44	4.6e+03	50000	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	58.00	83.91	1.1e+04	106746	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	57.00	130.60	26.48	363	
/cgml_9/gcgdata/geneseq/geneseq/NA198.DAT:AAV51465 +	57.00	122.18	77.97	965	
/cgml_9/gcgdata/geneseq/geneseq/NA199.DAT:AAV51465 +	57.00	116.41	163.44	1886	
/cgml_9/gcgdata/geneseq/geneseq/NA195.DAT:AAV51465 +	57.00	111.70	299.05	3259	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	57.00	89.34	5.3e+03	43676	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	57.00	78.03	2.2e+04	162450	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	126.37	45.53	456	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	125.31	52.19	516	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	119.51	109.70	1011	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	117.26	146.53	1314	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	117.26	146.53	1314	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	115.62	180.88	1590	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	114.54	207.57	1801	
/cgml_9/gcgdata/geneseq/geneseq/NA195.DAT:AAV51465 +	56.00	114.28	214.59	1856	
/cgml_9/gcgdata/geneseq/geneseq/NA198.DAT:AAV51465 +	56.00	113.83	227.39	1956	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	110.79	335.82	2784	
/cgml_9/gcgdata/geneseq/geneseq/NA197.DAT:AAV51465 +	56.00	99.44	1.4e+03	10401	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	98.34	1.7e+03	11820	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	77.17	2.5e+04	138169	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	76.95	2.6e+04	141589	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	76.95	2.6e+04	141589	
/cgml_9/gcgdata/geneseq/geneseq/NA200.DAT:AAV51465 +	56.00	76.95	2.6e+04	141589	

/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA35005 + 56.00 76.95 2.6e+04 141
/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21442 + 56.00 76.95 2.6e+04 141
/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21442 + 56.00 76.63 2.7e+04 146
/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21437 + 56.00 73.59 3.9e+04 209
/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAV89689 + 55.50 128.18 36.08 324

seq_name: /cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAV51465

seq_documentation_block:

ID: AAX51465 standard; cDNA: 421 BP.

XX AAX51465;

AC AAX51465;

DT 21-JUN-1999 (first entry)

XX Human secreted protein 5' EST SEQ ID NO:44.

DE Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX forensics; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductively hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

OS Homo sapiens.

XX

PN WO9906549-A2.

XX

XX 11-FEB-1999.

PD

XX 31-JUL-1998; 98WO-IB01231.

XX

XX 01-AUG-1997; 97US-0905279.

XX (GEST) GENSET.

FA

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI: 1999-153779/13.

DR P-PSDB; AAY12687.

XX

XX New nucleic acids encoding human secreted proteins - obtained from

PT cDNA libraries derived from testis, ovary, uterus and spleen tissue

PT

PS Claim 1; Page 177; 522pp; English.

XX

CC AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins, and encode the proteins given in AAY12681 to

CC AAY12913, respectively. The proteins given represent the signal peptide

CC and an N-terminal fragment of a secreted protein. The nucleic acid

CC sequences can be used for producing secreted human gene products. The

CC can also be used to develop products for diagnosis and therapy. The

CC proteins obtained may have cytokine activity, cell

CC proliferation/differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductively hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.

CC The sequences can also be used for obtaining corresponding promoter

CC sequences. The nucleic acids encoding the signal peptide can be used for

CC directing extracellular secretion of a polypeptide or the insertion of a

CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX SQ Sequence 421 BP; 96 A; 106 C; 110 G; 102 T; 7 other;

alignment_scores:

Quality: 119.00 Length: 23

Ratio: 5.174 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

```
US-09-215-435-179_COPY_1_23 x AAX51465
Align seg 1/1 to: AAX51465 from: 1 to: 421
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuPhePheLe 17
62 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTTCTCTCTCTTCTTCT 111
17 uPheLeuLeuThrArgGly 23
112 CTTCTCTCTCACCAGGGGC 130
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX97614
seq_documentation_block:
ID AAX97614 standard; DNA; 542 BP.
XX AC AAX97614;
XX DT
XX DE
XX DE Extended human secreted protein coding sequence, SEQ ID NO. 78.
XX KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
XX KW cellular differentiation; immune system regulator; anti-inflammatory;
XX KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
XX KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
XX KW genetic disease; ss.
XX OS Homo sapiens.
XX PN WO9931236-A2.
XX PD
XX PD
XX PD
XX PF 17-DEC-1998; 98WO-IB02122.
XX PR 10-AUG-1998; 98US-0096116.
XX PR 17-DEC-1997; 97US-0069957.
XX PR 09-FEB-1998; 98US-0074121.
XX PR 13-APR-1998; 98US-0081563.
XX PA (GEST ) GENSET.
XX PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX DR WPI; 1999-385906/32.
XX DR P-PSDB; AAY35930.
XX PT New isolated human secreted proteins
XX PS Claim 1; Page 212; 516pp; English.
XX CC This sequence represents an extended human secreted protein coding
XX CC sequence of the invention. The secreted proteins can be used in treating
XX CC or controlling a variety of human conditions. The secreted proteins may
XX CC act as cytokines or may affect cellular proliferation or differentiation
XX CC or may act as immune system regulators. haematopoiesis regulators, tissue
XX CC growth regulators, regulators of reproductive hormones or cell movement
XX CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
XX CC tumour inhibition activity. The DNAs can be used in forensic procedures
XX CC to identify individuals or in diagnostic procedures to identify
XX CC individuals having genetic diseases resulting from abnormal expression of
XX CC the genes corresponding to the extended cDNAs. They are also useful for
XX CC constructing a high resolution map of the human chromosomes. They can
XX CC also be used for gene therapy to control or treat genetic diseases.
XX SQ Sequence 542 BP; 132 A; 146 C; 126 G; 138 T; 0 other;
alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
```

```
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-179_COPY_1_23 x AAX97614
Align seg 1/1 to: AAX97614 from: 1 to: 542
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuPhePheLe 17
16 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTTCTCTCTTCTTCT 65
17 uPheLeuLeuThrArgGly 23
66 CTTCTCTCTCACCAGGGGC 84
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA26408
seq_documentation_block:
ID AAA26408 standard; cDNA; 680 BP.
XX AC AAA26408;
XX DT
XX DE 29-JUN-2000 (first entry)
XX DE Human secreted protein gene 63 SEQ ID NO:73.
XX KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX KW antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;
XX KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
XX KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
XX KW immune disease; inflammation; blood disorder; tumour; ss.
XX OS Homo sapiens.
XX PN WO200006598-A1.
XX PD
XX PD 10-FEB-2000.
XX PF 29-JUL-1999; 99WO-US17130.
XX PR 30-JUL-1998; 98US-0094657.
XX PR 05-AUG-1998; 98US-0095486.
XX PR 06-AUG-1998; 98US-0095454.
XX PR 06-AUG-1998; 98US-0095455.
XX PR 12-AUG-1998; 98US-0096319.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX DR WPI; 2000-195282/17.
XX DR P-PSDB; AAY91513.
XX PT New isolated human genes and the secreted polypeptides they encode,
XX PT useful for diagnosis and treatment of e.g. cancers, neurological
XX PT disorders, immune diseases, inflammation or blood disorders
XX PS Claim 1; Page 416; 634pp; English.
XX CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
XX CC proteins can have activities based on the tissues and cells they are
XX CC expressed in. Examples of the activities are: cytostatic;
XX CC immunosuppressive; antiHIV; antiinflammatory; nontropic; neuroprotective;
XX CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
XX CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
XX CC corresponding secreted proteins are useful for preventing, treating or
XX CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX CC pathological conditions can be diagnosed by determining the amount of the
XX CC proteins in a sample or by determining the presence of mutations in the
XX CC polynucleotides. Specific uses are described for each of the
```

CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AA26337 to AAA26345 and AA91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 680 BP; 163 A; 185 C; 167 G; 156 T; 9 other;

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x AAA26408 ..

Align seg 1/1 to: AAA26408 from: 1 to: 680

1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuLeuPhePheLeu 17
|||||
133 ATGATGCTACCCCAATGGCTGCTGCTGCTTCTCTCTCTCTTCT 182

17 upheLeuLeuThrArgGly 23
|||||
183 CTTCCTCTCACCAGGGGC 201

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF44847

seq_documentation_block:
ID AAF44847 standard; cDNA; 474 BP.

XX AAF44847;

XX 28-MAR-2001 (first entry)

XX Human breast cancer related protein coding sequence SEQ ID NO: 3.

XX Human; breast cancer; diagnosis; therapy; vaccine; ss.

XX Homo sapiens.

XX WO200078960-A2.

XX 28-DEC-2000.

XX 23-JUN-2000; 2000WO-US17536.

XX 23-JUN-1999; 99US-0140903.

XX 12-OCT-1999; 99US-0158980.

XX (CORI-) CORIXA CORP.

XX Yuqiu J, Mitcham JL;

XX WPI; 2001-041426/05.

XX New polynucleotides encoding breast tumor specific proteins, useful for
XX prevention, treatment and diagnosis of breast cancer -

XX Claim 5; Page 119; 165pp; English.

XX The present invention provides the coding sequences for a number of
XX breast cancer related proteins. These can be used in vaccinations

CC against, diagnosis of and treatment of cancer, particularly breast
CC cancer.
XX
SQ Sequence 474 BP; 191 A; 64 C; 123 G; 96 T; 0 other;

alignment_scores:
Quality: 62.50 Length: 19
Ratio: 3.676 Gaps: 1
Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
US-09-215-435-179_COPY_1_23 x AAF44847/rev ..

Align seg 1/1 to reverse of: AAF44847 from: 1 to: 474

2 MetLeuProGlnTrpLeuLeuLeuLeuPheLeuLeuPhePheLeuPh 18
:|||||
100 CTCTCTCTCT...CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 54

18 eLeuLeu 20
|||||
53 CCTCTTG 47

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA81713.

seq_documentation_block:
ID AAA81713 standard; DNA; 695 BP.

XX AAA81713;

XX 04-DEC-2000 (first entry)

XX N. meningitidis partial DNA sequence gnm_260 SEQ ID NO: 260.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX Meningococcus B; MenB; ds.

XX Neisseria meningitidis.

XX WO200022430-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

XX 09-OCT-1998; 98US-0103794.

XX 30-APR-1999; 99US-0132068.

XX (CHIR) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

XX Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

XX Rappuoli R, Pizza M;

XX WPI; 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea -

XX Claim 7; Page 1548; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. AAA81453 to AAA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
XX Neisseria DNA sequences and their corresponding proteins; AAA81254 to
XX AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
XX sequences, which are all used in the exemplification of the present

PS
XX AAQ67317 is the rat platelet factor 4 (PF4) promoter from Sprague-
CC Dawley strain. It was discovered that transgenic mice contg. a
CC nucleic acid construct encoding the SV40 early region tA58 mutant
CC regulated by this promoter unexpectedly develop varying levels of
CC thrombocytopenia and/or megakaryocyte leukaemia. These mice present
CC a novel and useful system for screening cpds. useful for treating
CC and/or preventing these diseases.
SO
XX Sequence 1264 BP; 324 A; 283 C; 276 G; 381 T; 0 other;

```

comment_block:
09-215-435-179_COPY_1_23 x AAQ67317 ..
- 09-215-435-179_COPY_1_23 x AAQ67317 from: 1 to: 1264

```

```

          4 proGlnTrpLeuLeuPheLeuPhePhePheLeuPheLeuLe 20
          ||||| |-----|:::|||||:::|||||:::|||||
          987 CCTCAGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1036
              20 u 20
                  |
                1037 G 1037

```

20 u 20
—
1037 G 1037

1037 G 1037
name: /cgnl_9/gcgcdata/geneseq/geneseqn/NA2000.DAT: AAC35324

name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT.AAC35324
documentation_block:
aac35324 standard: DNA: 705 BP.

```
documentation_block;
AAC35324 standard; DNA; 705 BP.
AAC35324;
AAC35324;
```

AAC35324;
17-OCT-2000 (first entry)
Arabidopsis thaliana DNA fragment SEQ ID NO: 9777.

17-OCT-2000 (first entry)
Arabidopsis thaliana DNA fragment SEQ ID NO: 9777.
Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; transcription; translation; promoter; termination sequence; ss.

Arabidopsis thaliana DNA fragment SEQ ID NO: 9777.

Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

Arabidopsis thaliana.

Arabidopsis thaliana.
EP1033405-A2.

EP1033405-A2.
06-SEP-2000.

06-SEP-2000.
25-FEB-2000: 2000EP-0301439.

25-FEB-2000; 2000EP-0301439.
25-FEB-1999; 99US-0121825.

25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.

05-MAR-1999; 99US-0123548.
09-MAR-1999; 99US-0125788.
23-MAR-1999; 99US-0126264.
25-MAR-1999; 99US-0126264.

23-MAR-1999; 99US-0126785.
25-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.

29-MAR-1999; 99US-0120703.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.

99US-0128234.
06-APR-1999;
99US-0128714.
08-APR-1999;
99US-0129845.
16-APR-1999;
99US-0130077.
18-APR-1999;

16-APR-1999; 99US-0129843.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
22-APR-1999; 99US-0130510.

21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
23-APR-1999; 99US-0131449.

23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
28-APR-1999; 99US-0132407.

30-APR-1999; 99US-0132046.
30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.

04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140333.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160770.

• 7

XX
07 JAN 1977

XX	Key	Location/Qualifiers
FH	CDS	1386..3119
FT		


```
FT      /*tag= a
FT      /product= thrombomodulin
PN      JP09268200-A.
PN      14-OCT-1997.
XX      01-APR-1996; 96JP-0078494.
XX      01-APR-1996; 96JP-0078494.
XX      (ASAH ) ASahi KASEI KOGYO KK.
XX      WPI; 1997-554712/51.
XX      P-PSDB; AAW30845.
XX      Recombinant production of rat thrombomodulin - which promotes
XX      thrombin activation of protein C
XX      Disclosure; Pages 22-25; 28pp; Japanese.
XX      This DNA encodes a rat thrombomodulin, which upon binding thrombin,
XX      promotes thrombin activation of protein C. Host cells transformed with
XX      the DNA can be used for the large scale recombinant production of the
XX      peptide, or rat thrombomodulin.
XX      Sequence 5283 BP; 1244 A; 1324 C; 1330 G; 1385 T; 0 other;

alignment_scores:
  Quality: 59.00      Length: 17
  Ratio: 3.933        Gaps: 0
  Percent Similarity: 88.235 Percent Identity: 64.706

alignment_block:
US-09-215-435-179_COPY_1_23 x AAT91972 ..
Align seg 1/1 to: AAT91972 from: 1 to: 5283
6 TrpLeuLeuLeuPheLeuPhePheLeuPheLeuLeuThrAr 22
||||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3874 TGGGTTTATGATTGCTGTTGTTGTTTTTTTTTTATTATTAGATGACAAA 3923
22 g 22
3924 A 3924

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT.AAV90279
seq_documentation_block:
ID      AAV90279 standard; cDNA; 303 BP.
XX
AC      AAV90279;
XX
XX      15-FEB-1999 (first entry)
XX
XX      EST clone D1387.
XX
XX      Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
XX      tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
XX      receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
XX      gene therapy; ss.
XX
XX      Homo sapiens.
XX
XX      WP9845436-A2.
XX
XX      15-OCT-1998.
XX
XX      10-APR-1998; 98WO-US06955.
XX
XX      10-APR-1997; 97US-0838821.
XX
XX
```

```
PA      (GEMY ) GENETICS INST INC.
XX
XX      Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
XX      Racie LA, Spaulding V, Treacy M;
XX      WPI; 1999-070077/06.
XX
XX      New polynucleotides encoding human secreted proteins - derived from
XX      e.g. human blood, kidney, foetal lung, placenta, testes, brain,
XX      ovary, pituitary, retina and colon cDNA libraries.
XX
XX      Claim 1; Page 489; 618pp; English.
XX
XX      The present sequence represents a human expressed sequence tag (EST).
XX      The polynucleotide, which is a secreted EST, and the encoded protein
XX      are predicted to have useful biological activities which would make
XX      them suitable for treating, preventing or ameliorating medical
XX      conditions in humans and animals, although no supporting data is
XX      given. Suggested activities include nutritional activity, immune
XX      stimulating or suppressing activity, haematopoiesis regulating
XX      activity, tissue growth activity, activin/inhibin activity,
XX      chemotactic/chemokinetic activity, haemostatic and thrombolytic
XX      activity, receptor/ligand activity, anti-inflammatory activity,
XX      cadherin/tumour invasion suppressor activity, tumour inhibition
XX      activity. The polynucleotide may also be useful for gene therapy.
XX
XX      Sequence 303 BP; 121 A; 21 C; 124 G; 37 T; 0 other;

alignment_scores:
  Quality: 58.00      Length: 15
  Ratio: 4.462        Gaps: 0
  Percent Similarity: 86.667 Percent Identity: 73.333

alignment_block:
US-09-215-435-179_COPY_1_23 x AAV90279/rev ..
Align seg 1/1 to reverse of: AAV90279 from: 1 to: 303
4 ProGlnTrpLeuLeuLeuPheLeuPheLeuPhePheLeuPhe 18
||||| |||:||||| |||:||||| |||:||||| |||:|||||
216 CCAACTGGTGCCTTTTGTATTCTCTCTACTCTCTCTCTTC 172

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1996.DAT.AAT05868
seq_documentation_block:
ID      AAT05868 standard; DNA; 3399 BP.
XX
AC      AAT05868;
XX
XX      14-AUG-1996 (first entry)
XX
XX      Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.
XX
XX      Chicken leucocytozoan; immunogen; recombinant vaccine; protection;
XX      immunisation; vaccination; ss.
XX
XX      Chicken leucocytozoan.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..3399
XX              /*tag= a
XX      misc_feature      1150..3218
XX              /*tag= b
XX              /note= "fragment referred to in the claims, for
XX              use as insert in a recombinant vaccine
XX              against chicken leucocytozoan disease"
XX
XX      JP07284392-A.
XX
XX      31-OCT-1995.
XX
XX      19-APR-1994; 94JP-0080643.
XX
```

Thu Nov 15 10:52:03 2001

```

XX 19-APR-1994; 94JP-0080643.
XX (DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
PA (KITA ) KITASATO KENKYUSHO SH.
XX WPI; 1996-006311/01.
XX P-PSDB; AAR97866.
XX Chicken leucocytozoan immunogenic protein - used in a recombinant
PT vaccine against chicken leucocytozoan disease
XX Claim 6; Page 6-9; 35pp; Japanese.
XX AAT05868 encodes a chicken leucocytozoan immunogenic protein, this DNA
CC or a fragment of it can be used in a recombinant vaccine to immunise
CC against chicken leucocytozoan disease. The DNA is used in a vector
CC and operatively linked to an expression regulatory sequence as in
CC standard practice.
XX SQ Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 other;

alignment_scores:
    Quality: 58.00 Length: 20
    Ratio: 3.222 Gaps: 1
Percent Similarity: 90.000 Percent Identity: 65.000

alignment_block:
US-09-215-435-179_COPY_1_23 x AAT05868/rev ..

Align seg 1/1 to reverse of: AAT05868 from: 1 to: 3399

1 MetMetLeuProGlnTrpLeuLeuLeuLeu.PheLeuLeuPhePheL 17
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
2869 CTCTCTCTCTTATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2820

17 eupheLeu 19
   ::::::::::
2819 TCTTCCTC 2812

```

; CITY: Thousand Oaks

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..2848
PCT-US95-03747-1

alignment_scores:
  Quality: 57.00      Length: 17
  Ratio: 3.800       Gaps: 0
  Percent Similarity: 88.235  Percent Identity: 70.588

alignment_block:
  US-09-215-435-179_COPY_1_23 x PCT-US95-03747-1/rev ..
  Align seg 1/1 to reverse of: PCT-US95-03747-1 from: 1 to: 3259

      4  proGlnTrpLeuLeuLeuPheLeuLeuPhePheLeuLeuLe 20
      ||||| |||:||||| :|||||:|||||:|||||:|||||
      1551 CCACAGGCGCTCATCTCCACGCTCCTCTTCTTCTTCTTCT 1502

      20 u 20
      |
      1501 c 1501

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-356-952-12

seq_documentation_block:
; Sequence 12, Application US/09356952
; Patent No. 6117863
; GENERAL INFORMATION:
; APPLICANT: Boriack-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Sogli, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; CURRENT FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 43676
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-356-952-12

alignment_scores:
  Quality: 57.00      Length: 15
  Ratio: 3.800       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 80.000

alignment_block:
  US-09-215-435-179_COPY_1_23 x US-09-356-952-12/rev ..
  Align seg 1/1 to reverse of: US-09-356-952-12 from: 1 to: 43676

      6  TrpLeuLeuLeuPheLeuLeuPhePheLeuPheLeuLeu 20
      :|||:|||||:|||||:|||||:|||||:|||||:|||||
      20963 TTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20919

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-464-523B-12

seq_documentation_block:
; Sequence 12, Application US/08464523B
; Patent No. 5723761

```

```

;
; GENERAL INFORMATION:
; APPLICANT: Toni A. Voelker
; APPLICANT: Ling Yuan
; APPLICANT: Jean Kridl
; APPLICANT: Deborah Hawkins
; APPLICANT: Aubrey Jones
; TITLE OF INVENTION: Plant Acyl ACP Thioesterase
; TITLE OF INVENTION: Sequences
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,523B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13131
; FILING DATE: 10-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,004
; FILING DATE: 10-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 100-1W0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1856 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-464-523B-12

alignment_scores:
  Quality: 56.00      Length: 15
  Ratio: 4.000        Gaps: 0
  Percent Similarity: 93.333  Percent Identity: 73.333

alignment_block:
US-09-215-435-179_COPY_1_23 x US-08-464-523B-12  ..
Align seg 1/1 to: US-08-464-523B-12 from: 1 to: 1856

      5 GlnTrpLeuLeuLeuPheLeuPheLeuPhePheLeu 19
      144 AGATGCTTCTACTGCTGCTACTGCTGCTTTTTCACGATTCTT 188

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-420-443-5

seq_documentation_block:
; Sequence 5, Application US/08420443
; Patent No. 5607834
; GENERAL INFORMATION:
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: MIYAKE, Kazue
; APPLICANT: FUJITA, Tomoko
; APPLICANT: SATO, Yoshimi
; APPLICANT: UEMORI, Takashi
; GENERAL INFORMATION:
; Patent No. 6218150
; Sequence 63, Application US/09446504
; seq_documentation_block:
; seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-446-504-63
```

```

;
; GENERAL INFORMATION:
; APPLICANT: C. Bruce Bagwell
; TITLE OF INVENTION: NUCLEIC ACID PROBES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/420,443
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/990,298
; FILING DATE: 19921210
; ATTORNEY/AGENT INFORMATION:
; NAME: Y. Rocky Tsao
; REGISTRATION NUMBER: 34,053
; REFERENCE/DOCKET NUMBER: 05663/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: N stands for a 6-carbon
; OTHER INFORMATION: molecular linker. See Fig. 7
; OTHER INFORMATION: of the specification
; US-08-420-443-5

alignment_scores:
  Quality: 55.00      Length: 18
  Ratio: 3.235        Gaps: 0
  Percent Similarity: 94.444  Percent Identity: 50.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-08-420-443-5  ..
Align seg 1/1 to: US-08-420-443-5 from: 1 to: 83

      5 GlnTrpLeuLeuLeuPheLeuPhePheLeuPheLeuLeu 21
      16 AAGTGGTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 65

      21 rArg 22
      :
      66 GAAA 69

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-446-504-63

seq_documentation_block:
; Sequence 63, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazue
; APPLICANT: MUKAI, Hiroyuki
```

APPL
FILING
ATTORNEY

Align seg 1/1 to reverse of: US-09-115-446-1 from: 1 to: 2521

6 TrpLeuLeuLeuPheLeuPhePheLeuPheLeuThrAr 22
|||||
412 TGCGTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTATCCCCAG 363

22 ggly 23
||||
362 AGGA 359

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-115-446-5

seq_documentation_block:

; Sequence 5, Application US/09115446

; Patent No. 6165719

; GENERAL INFORMATION:

; APPLICANT: Chandy, George K.

; APPLICANT: Gargus, Jay J.

; APPLICANT: Gutman, George

; APPLICANT: Fantino, Emmanuelle

; APPLICANT: Kalman, Katarin

; TITLE OF INVENTION: hKCA3/KCN3 SMALL CONDUCTANCE CALCIUM

; TITLE OF INVENTION: ACTIVATED POTASSIUM CHANNEL: A DIAGNOSTIC

; TITLE OF INVENTION: MARKER AND THERAPEUTIC TARGET

; FILE REFERENCE: 07306/014001

; CURRENT APPLICATION NUMBER: US/09/115,446

; CURRENT FILING DATE: 1998-07-14

; EARLIER APPLICATION NUMBER: 60/052,556

; EARLIER FILING DATE: 1997-07-15

; EARLIER APPLICATION NUMBER: 60/070,741

; EARLIER FILING DATE: 1998-01-08

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 2526

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-115-446-5

alignment_scores:

Quality: 55.00 Length: 18

Ratio: 3.235 Gaps: 0

Percent Similarity: 94.444 Percent Identity: 61.111

alignment_block:

US-09-215-435-179_COPY_1_23 x US-09-115-446-5/rev ..

Align seg 1/1 to reverse of: US-09-115-446-5 from: 1 to: 2526

6 TrpLeuLeuLeuPheLeuPhePheLeuPheLeuThrAr 22
|||||
417 TGCGTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTATCCCCAG 368

22 ggly 23
||||

367 AGGA 364

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-446-504-83

seq_documentation_block:

; Sequence 83, Application US/09446504

; Patent No. 6218150

; GENERAL INFORMATION:

; APPLICANT: UEMORI, Takashi

; APPLICANT: SATO, Yoshimi

; APPLICANT: FUJITA, Tomoko

; APPLICANT: MIYAKE, Kazuo

; APPLICANT: MUKAI, Hiroyuki

; APPLICANT: ASADA, Kiyozo

; APPLICANT: KATO, Ikunoshin

; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS

; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/446,504
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3574
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-446-504-83

alignment_scores:

Quality: 55.00 Length: 13

Ratio: 4.231 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 76.923

alignment_block:

US-09-215-435-179_COPY_1_23 x US-09-446-504-83/rev ..

Align seg 1/1 to reverse of: US-09-446-504-83 from: 1 to: 3574

6 TrpLeuLeuLeuPheLeuPhePheLeuPheLeuPhe 18
|||||
1398 TGGCTTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTC 1360

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-431-080-27

seq_documentation_block:

; Sequence 27, Application US/08431080

; Patent No. 5698686

; GENERAL INFORMATION:

; APPLICANT: Gottschling, Daniel E.

; APPLICANT: Singer, Miriam S.

; TITLE OF INVENTION: Telomerase Compositions and Methods

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TEXAS

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/431,080

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: SN 08/326,781

; FILING DATE: October 20, 1994

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, David L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: ARCD-155/PAR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TELEX: 79-0924

; INFORMATION FOR SEQ ID NO: 27:

۱۰۰


```

alignment_block:
US-09-215-435-179_COPY_1_23 x US-08-476-866-20 ..
Align seg 1/1 to: US-08-476-866-20 from: 1 to: 10079

7 LeuLeuLeuPheLeuLeuPhePheLeuPheLeuThrArgG1 23
|||||.....:|||||||:|||||:|||||:|||||
1992 CACCCTCCTCTCTTTTTCCTTTTTTTTTTTTTTTTGTGACATGG 2041

23 Y 23
      |
2042 T 2042

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-009-913-287

seq_documentation_block:
; Sequence 287, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axys Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 287:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-009-913-287

alignment_scores:
Quality: 54.00 Length: 15
Ratio: 3.857 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 73.333

alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-009-913-287/rev ..
Align seg 1/1 to reverse of: US-09-009-913-287 from: 1 to: 8

6 TrpLeuLeuLeuPheLeuLeuPhePheLeuPheLeu 20
|||||.....:|||||||:|||||:|||||:|||||
63 TTTCCTTTCTTCTTCTACTTCTTCTTCTTCTTCTTCTTCTT 19

```

THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-179_COPY_1_23 to: Pending_Patents_NA_Main: * out_format : pfs
Date: Nov 15, 2001 3:55 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=framet_p2n.model -DEV=xlp
-O=cgml_1/USPTO.spool/US09215435/runat_14112001_141105_2173/app_query.fasta_1.1519
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnmp
-GAPOP=12.000 -GAPEXT=4.500 -MINMATCH=0.100 -LOOPCL=0.000
-XGAPEXT=0.500 -FGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blosum62 -TRANS=human40 cdi -LIST=45 -DOCALLIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435@cgnl_1_12482 -NCPU=6 -ICPU=3 -LONGLOG
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-215-435-179_COPY_1_23
Query length: 23
Database: Pending_Patents_NA_Main: *
Database sequences: 17159718
Database length: 173266264
Search time (sec): 11750.060000

score_list:	Strd Orig	zScore	Escore	len	Documentation	
/cgml_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-44 +	119.00	236.19	0.0008	421		
/cgml_7/ptodata/1/pna/US097A_COMB.seq:US-09-731-872-89 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-5913 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US098_COMB.seq:US-09-876-997-89 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US6014_COMB.seq:US-60-147-499-1374 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US6016_COMB.seq:US-60-169-629-108 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US6018_COMB.seq:US-60-187-470-108 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-5913 +	119.00	234.53	0.0010	514		
/cgml_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-4817 +	119.00	234.23	0.0011	533		
/cgml_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-4817 +	119.00	234.23	0.0011	533		
/cgml_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-78 +	119.00	234.09	0.0011	542		
/cgml_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-78 +	119.00	233.74	0.0011	565		
/cgml_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-78 +	119.00	232.20	0.0014	680		
/cgml_7/ptodata/1/pna/US094_COMB.seq:US-09-489-847-73 +	119.00	232.20	0.0014	680		
/cgml_7/ptodata/1/pna/US6021_COMB.seq:US-60-212-664-192 +	119.00	186.18	0.5005	171		
/cgml_7/ptodata/1/pna/US6021_COMB.seq:US-60-212-664-192 +	119.00	186.18	0.5005	171		
/cgml_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481-520 +	70.00	139.89	189.44	488		
/cgml_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481-520 +	70.00	139.89	189.44	488		
/cgml_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481-520 +	70.00	139.89	189.44	488		
/cgml_7/ptodata/1/pna/US092_COMB.seq:US-09-244-000A-84931 +	69.00	143.55	118.52	2		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-242-679-659 +	67.00	101.89	2.5e+04	233		
/cgml_7/ptodata/1/pna/US6023_COMB.seq:US-60-230-435-750 +	67.00	99.46	3.4e+04	313		
/cgml_7/ptodata/1/pna/US6021_COMB.seq:US-60-213-173-34 -	67.00	99.08	3.6e+04	3278		
/cgml_7/ptodata/1/pna/US6021_COMB.seq:US-60-213-173-34 -	67.00	99.08	3.6e+04	3278		
/cgml_7/ptodata/1/pna/US096D_COMB.seq:US-09-666-355A-7255 -	66.00	133.15	450.06	4		
/cgml_7/ptodata/1/pna/US096D_COMB.seq:US-09-666-355A-7255 -	66.00	132.08	515.91	491		
/cgml_7/ptodata/1/pna/US098_COMB.seq:US-09-864-761-931 +	66.00	132.08	515.91	491		
/cgml_7/ptodata/1/pna/US6023_COMB.seq:US-60-236-359-1002 +	66.00	132.08	515.91	491		
/cgml_7/ptodata/1/pna/US096E_COMB.seq:US-09-654-617-197755 -	66.00	129.19	747.41	49		
/cgml_7/ptodata/1/pna/US096E_COMB.seq:US-09-654-617-197755 -	66.00	129.19	747.41	49		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-248-505-1523 +	66.00	123.53	1.5e+03	13		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-248-505-1523 -	66.00	121.88	1.9e+03	16		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-248-505-1523 -	66.00	117.19	3.5e+03	294		
/cgml_7/ptodata/1/pna/US099_COMB.seq:US-09-950-083-7308 +	66.00	103.85	1.9e+04	148		
/cgml_7/ptodata/1/pna/US099_COMB.seq:US-09-950-083-7308 +	66.00	101.14	2.7e+04	20		
/cgml_7/ptodata/1/pna/US099_COMB.seq:US-09-950-083-7308 +	66.00	101.14	2.7e+04	202		
/cgml_7/ptodata/1/pna/US097C_COMB.seq:US-09-764-855-173 +	66.00	101.14	2.7e+04	202		
/cgml_7/ptodata/1/pna/US099_COMB.seq:US-09-950-083-5384 +	66.00	101.14	2.7e+04	202		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-248-505-191 +	66.00	95.56	5.6e+04	396		
/cgml_7/ptodata/1/pna/US6024_COMB.seq:US-60-248-505-191 +	66.00	95.56	5.6e+04	396		

seq_name: /cgml_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-44
seq documentation block:
; Sequence 44. Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF INVENTIONS: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,279
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Testis
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 62..130
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: seq FLLFFLLTRG/SL
; OTHER INFORMATION: seq FLLFFLLTRG/SL
US-08-905-279-44
alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-179_COPY_1_23 x US-08-905-279-44
Align seg 1/1 to: US-08-905-279-44 from: 1 to: 421
1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLeu 17
62 ATGATGCTACCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTTCTTCT 111
17 uPheLeuLeuThrArgGly 23
|||||

!

```
Align seg 1/1 to: US-09-876-997-89 from: 1 to: 514
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLe 17
99 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTCTCTCTCTCTCT 148
17 uPheLeuLeuThrArgGly 23
149 CTTCTCTCTCACCAGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/US6014_COMB.seq:US-60-147-499-1374
seq_documentation_block:
; Sequence 1374, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1374
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..356
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.80000019073486
; OTHER INFORMATION: seq FLLFFFLLRG/SL
US-60-147-499-1374

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-60-147-499-1374 ..
Align seg 1/1 to: US-60-147-499-1374 from: 1 to: 514
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLe 17
99 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTCTCTCTCTCTCT 148
17 uPheLeuLeuThrArgGly 23
149 CTTCTCTCTCACCAGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-169-629-108
seq_documentation_block:
; Sequence 108, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 108
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 113..364
; NAME/KEY: sig_peptide
; LOCATION: 113..172
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.3
; OTHER INFORMATION: seq SLLSLPPHQLT/FS
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 478..483
; NAME/KEY: polyA_site
; LOCATION: 500..514
US-60-169-629-108

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-60-169-629-108 ..
Align seg 1/1 to: US-60-169-629-108 from: 1 to: 514
1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLe 17
99 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTCTCTCTCTCTCT 148
17 uPheLeuLeuThrArgGly 23
149 CTTCTCTCTCACCAGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-187-470-108
seq_documentation_block:
; Sequence 108, Application US/60187470
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: 78.US2.PRO
; CURRENT APPLICATION NUMBER: US/60/187,470
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 108
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 113..364
; NAME/KEY: sig_peptide
; LOCATION: 113..172
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.3
; OTHER INFORMATION: seq SLLSLPPHQLT/FS
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 478..483
; NAME/KEY: polyA_site
; LOCATION: 500..514
```

```

; sequence 4817, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..461
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; score 9.80000019073486
; OTHER INFORMATION: seq FLFFFLFLTRG/SL
US-09-834-366-4817

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-834-366-4817 ..
Align seg 1/1 to: US-09-834-366-4817 from: 1 to: 533

1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPheLeu 17
|||||
99 ATGATGCTACCCCAATGGCTGCTGTCTTCTTCCTTTCTTCT 148

17 uPheLeuLeuThrArgGly 23
|||||
149 CTTCCTCCTCACCGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/us6019_COMB.seq.us-60-197-873-4817

seq_documentation_block:
; Sequence 4817, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 113..364
; NAME/KEY: sig_peptide
; LOCATION: 113..172
; OTHER INFORMATION: Von Heijne matrix
; score 4.30000019073486
; OTHER INFORMATION: seq SLLLSLPPHQLT/FS
US-60-197-873-5913

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-60-197-873-5913 ..
Align seg 1/1 to: US-60-197-873-5913 from: 1 to: 514

1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPheLeu 17
|||||
99 ATGATGCTACCCCAATGGCTGCTGTCTTCTTCCTTTCTTCT 148

17 uPheLeuLeuThrArgGly 23
|||||
149 CTTCCTCCTCACCGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/us098_COMB.seq.us-09-834-366-4817

seq_documentation_block:

```

OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 9.8000019073486
OTHER INFORMATION: seq FLLFFFLLLTRG/SL
US-60-197-873-4817

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_copy_1_23 x US-60-197-873-4817 ..

Align seg 1/1 to: US-60-197-873-4817 from: 1 to: 533

1 MetMetLeuProGlnTrpLeuLeuLeuLeuLeuLeuLeuLeuPhePhePheLe 17
|||||
99 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCTTCTTCT 148
17 uPheLeuLeuThrArgGly 23
|||||
149 CTTCCTCTCACCAGGGC 167

seq_name: /cgnl_7/ptodata/1/pna/US092_COMB.seq:US-09-215-435-78

seq_documentation_block:
Sequence 78, Application US/09215435A

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert Aymeric,
TITLE OF INVENTION: Bouqueleret Lydie
FILE REFERENCE: GENSET 019A
CURRENT APPLICATION NUMBER: US/09/215,435A
CURRENT FILING DATE: 1998-12-17
EARLIER APPLICATION NUMBER: 60/069,957
EARLIER FILING DATE: 1997-12-17
EARLIER APPLICATION NUMBER: 60/074,121
EARLIER FILING DATE: 1998-2-9
EARLIER APPLICATION NUMBER: 60/081,563
EARLIER FILING DATE: 1998-4-13
EARLIER APPLICATION NUMBER: 60/096,116
EARLIER FILING DATE: 1998-8-10
NUMBER OF SEQ ID NOS: 519
SOFTWARE: Patent.pm

SEQ ID NO 78
LENGTH: 542
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 16..378
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 16..84
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 9.8
OTHER INFORMATION: seq FLLFFFLLLTRG/SL

FEATURE:
NAME/KEY: polyA_signal
LOCATION: 502..507
FEATURE:
NAME/KEY: polyA_site
LOCATION: 522..542
US-09-215-435-78

alignment_scores:
Quality: 119.00 Length: 23
Ratio: 5.174 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_copy_1_23 x US-09-215-435-78 ..
Align seg 1/1 to: US-09-215-435-78 from: 1 to: 542
1 MetMetLeuProGlnTrpLeuLeuLeuLeuLeuLeuLeuLeuPhePhePheLe 17
|||||
16 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCTTCTTCT 65
17 uPheLeuLeuThrArgGly 23
|||||
66 CTTCCTCTCACCAGGGC 84
seq_name: /cgnl_7/ptodata/1/pna/US6009_COMB.seq:US-60-096-116-61
seq_documentation_block:
Sequence 61, Application US/60096116
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste Dumas Milne
APPLICANT: Duclert, Aymeric
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SECRETED PROTEINS
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/096,116
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET.038PR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 565 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Testis
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 16..84
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 9.8
OTHER INFORMATION: seq FLLFFFLLLTRG/SL
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 502..507
FEATURE:
NAME/KEY: poly_a
LOCATION: 529..542
IDENTIFICATION METHOD: blastn2
FEATURE:
NAME/KEY: vrt
LOCATION: complement(33..74)
IDENTIFICATION METHOD: fasta
OTHER INFORMATION: identity 88

Thu Nov 15 10:52:03 2001

us-09-215-435-179_copy_1_23.rnrm

```

; OTHER INFORMATION: region 21722..21764
; FEATURE:
; NAME/KEY: vit
; LOCATION: 511..545
; IDENTIFICATION METHOD: fasta
; OTHER INFORMATION: identity 83
; OTHER INFORMATION: region 4627..4663
; OTHER INFORMATION: id AF048728
US-60-096-116-61

alignment_scores:
  Quality: 119.00      Length: 23
  Ratio: 5.174        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x US-60-096-116-61 ..
Align seg 1/1 to: US-60-096-116-61 from: 1 to: 565
1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLe 17
|||||
16 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTGCTGCTTCTTCTTCT 65

17 uPheLeuLeuThrArgGly 23
|||||
66 CTTCTCTCTCACCAGGGGC 84

seq_name: /cgnl_7/ptodata/1/pna/PCTUS_COMB.seq:PCT-US99-17130-73

seq_documentation_block:
; Sequence 73, Application PCT/US9917130
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/17130
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
US-09-489-847-73

; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
PCT-US99-17130-73

alignment_scores:
  Quality: 119.00      Length: 23
  Ratio: 5.174        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_1_23 x PCT-US99-17130-73 ..
Align seg 1/1 to: PCT-US99-17130-73 from: 1 to: 680
1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuPhePheLe 17
|||||
133 ATGATGCTACCCCAATGGCTGCTGCTGCTGCTGCTGCTTCTTCTTCT 182

17 uPheLeuLeuThrArgGly 23
|||||
183 CTTCTCTCTCACCAGGGGC 201

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-489-847-73

seq_documentation_block:
; Sequence 73, Application US/09489847
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
US-09-489-847-73

```



```

alignment_scores:
  Quality: 119.00      Length: 23
  Ratio: 5.174         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-179_COPY_1_23 x US-09-489-847-73  ..
  Align seg 1/1 to: US-09-489-847-73 from: 1 to: 680

1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuLeuPhePheLe 17
|||||
133 ATGATGCTACCCCAATGGCTGCTGCTGTGCTTCTCTCTCTTCT 182

17 uPheLeuLeuThrArgGly 23
|||||
183 CTTCCTCTCTACCCAGGGC 201

seq_name: /cgnl_7/ptodata/1/pna/US6021_COMB.seq:US-60-212-664-192

seq_documentation_block:
; Sequence 192, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Spier, Gene
; APPLICANT: Greenberg, Simon
; APPLICANT: Rabkin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000687
; CURRENT APPLICATION NUMBER: US/60/212,664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 171973
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(171973)
; OTHER INFORMATION: n = A,T,C or G
US-60-212-664-192

alignment_scores:
  Quality: 119.00      Length: 23
  Ratio: 5.174         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-179_COPY_1_23 x US-60-212-664-192  ..
  Align seg 1/1 to: US-60-212-664-192 from: 1 to: 171973

1 MetMetLeuProGlnTrpLeuLeuLeuLeuPheLeuLeuPhePheLe 17
|||||
83353 ATGATGCTACCCCAATGGCTGCTGCTGTGCTTCTCTCTTCTTCT 83402

17 uPheLeuLeuThrArgGly 23
|||||
83403 CTTCCTCTCTACCCAGGGC 83421

```

THIS PAGE BLANK (USPTO)


```

; PRIOR APPLICATION NUMBER: 60/055,312
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,807
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,386
; PRIOR FILING DATE: 1997-08-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-730-28

alignment_scores:
    Quality: 55.50      Length: 22
    Ratio: 2.775       Gaps: 1
    Percent similarity: 90.909    Percent Identity: 54.545

alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-969-730-28/rev ..
Align seg 1/1 to reverse of: US-09-969-730-28 from: 1 to: 794

3 LeuProGlnTrpLeuLeuLeuPheLeuPheLeuPheLeuPheLe 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
579 ATGCCAGCTCATCTGTTATTTTATTTATTTATTTATTTACTT 530

19 uLeu...ThrArgGly 23
||||| |||||||
529 GCTATGGACTAGGGGA 514

seq_name: /cgml_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-108-010A-11

seq_documentation_block:
; Sequence 11, Application US/09108010A
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; TITLE OF INVENTION: CLASSES OF SOYBEAN SEED
; TITLE OF INVENTION: PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,010A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481

```

```

; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-108-010A-11

alignment_scores:
Quality: 55.50 Length: 24
Ratio: 2.921 Gaps: 1
Percent Similarity: 79.167 Percent Identity: 54.167

alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-108-010A-11/rev ..
Align seg 1/1 to reverse of: US-09-108-010A-11 from: 1 to: 1488

3 LeuProGlnTrpLeuLeuPheLeuPhePhe..... 16
||| |||||||:|||||:|||||:|||||:|||||
873 CTGCACGTGGCTCATCCTCTTCTTCCTTCCTTCCTCGGGTC 824

17 .LeuPheLeuLeuThrArgGly 23
|||:|||||:|||||
823 TTTGTTGCTGCTCGTCGCGGG 802

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-10083
seq_documentation_block:
; Sequence 10083, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10083
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(426)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-10083

alignment_scores:
Quality: 54.50 Length: 23
Ratio: 3.028 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-922-340-10083/rev ..
Align seg 1/1 to reverse of: US-09-922-340-10083 from: 1 to: 426

2 MetLeuProGlnTrpLeu.....LeuLeuPheLeuLeuPhePh 15
::||| ::||| ||| ::|||:|||||:|||||:
346 GTCTCCACGCGTGCACTCAGTCATCTCTTCCTTCCTTCCT 297

15 ePheLeuPheLeuThr 21
:::|||||:|||||

```

•

```

alignment_scores:
  Quality: 54.00      Length: 13
  Ratio: 4.154        Gaps: 0
  Percent Similarity: 100.000    Percent Identity: 69.231

alignment_block:
  US-09-215-435-179_COPY_1_23 x US-09-973-278-792
  ..

```

```
alignment_scores:
  Quality: 53.00
  Length: 16
```

```
; ORGANISM: Homo sapiens
US-09-886-055-236

alignment_scores:
  Quality: 53.00      Length: 18
  Ratio: 3.533       Gaps: 0
  Percent Similarity: 83.333   Percent Identity: 44.444

alignment_block:
US-09-215-435-179_COPY_l_23 x US-09-886-055-236 ..
Align seg 1/1 to: US-09-886-055-236 from: 1 to: 945

4 ProGlnTrpLeuLeuPheLeuPheLeuPheLeuPheLeuLeu 20
||||| ||||||| :||| |||||||:||||| :|||:|||||
76 CCAGAGTGGAAATGCCCTGCTCTTGTGCTGCTATCTCAT 125

20 uThr 21
126 CACT 129

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-886-055-414

seq_documentation_block:
; Sequence 414, Application US/09886055
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; PRIOR FILING DATE: 2001-06-22
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 414
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-414

alignment_scores:
  Quality: 53.00      Length: 18
  Ratio: 3.786       Gaps: 0
  Percent Similarity: 77.778   Percent Identity: 61.111

alignment_block:
US-09-215-435-179_COPY_l_23 x US-09-886-055-414 ..
Align seg 1/1 to: US-09-886-055-414 from: 1 to: 954

4 ProGlnTrpLeuLeuPheLeuPheLeuPheLeuPheLeuLeu 20
||||| |||||||:|||||:|||||:|||||:|||||
61 CCTCCCTCCAGCTGCTCTCTTGTGCTTTTGTGCAATTACCTTCT 110

20 uThr 21
111 GACA 114

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-969-730-67

seq_documentation_block:
; Sequence 67, Application US/09969730
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P201392
; CURRENT APPLICATION NUMBER: US/09/969,730
; CURRENT FILING DATE: 2001-10-04
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-886-055-236

alignment_scores:
  Quality: 53.00      Length: 16
  Ratio: 4.417       Gaps: 0
  Percent Similarity: 75.000   Percent Identity: 68.750

alignment_block:
US-09-215-435-179_COPY_l_23 x US-09-922-340-2910 ..
Align seg 1/1 to: US-09-922-340-2910 from: 1 to: 424

5 GlnTrpLeuLeuPheLeuPheLeuPheLeuPheLeuLeu 20
||||| ||||||| :||| |||||||:||||| |||
110 CAGTGTCTCTGCTGCTGCTGCTGCTGCTCTTCTTCTTCGCTTG 157

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-2909

seq_documentation_block:
; Sequence 2909, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2909
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(477)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-2909

alignment_scores:
  Quality: 53.00      Length: 16
  Ratio: 4.417       Gaps: 0
  Percent Similarity: 75.000   Percent Identity: 68.750

alignment_block:
US-09-215-435-179_COPY_l_23 x US-09-922-340-2909 ..
Align seg 1/1 to: US-09-922-340-2909 from: 1 to: 477

5 GlnTrpLeuLeuPheLeuPheLeuPheLeuPheLeuLeu 20
||||| ||||||| :||| |||||||:||||| |||
110 CAGTGTCTCTGCTGCTGCTGCTGCTGCTCTTCTTCTTCGCTTG 157

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-886-055-236

seq_documentation_block:
; Sequence 236, Application US/09886055
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 236
; LENGTH: 945
; TYPE: DNA
```


;; PRIOR APPLICATION NUMBER: 09/774,639
;; PRIOR FILING DATE: 2001-02-01
;; PRIOR APPLICATION NUMBER: 60/238,291
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 09/244,112
;; PRIOR FILING DATE: 1999-02-04
;; PRIOR APPLICATION NUMBER: PCT/US98/16235
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/056,371
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,732
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,366
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,364
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,370
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,367
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,365
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,731
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,557
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/056,563
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 60/055,970
;; PRIOR FILING DATE: 1997-08-18
;; PRIOR APPLICATION NUMBER: 60/055,986
;; PRIOR FILING DATE: 1997-08-18
;; PRIOR APPLICATION NUMBER: 60/054,803
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/054,804
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/054,809
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/054,806
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/055,310
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/054,798
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/055,309
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/055,312
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/054,807
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: 60/055,386
;; PRIOR FILING DATE: 1997-08-05
;; NUMBER OF SEQ ID NOS: 373
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 67
;; LENGTH: 1323
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (1086)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (1087)
;; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-730-67

Quality: 53.00 Length: 17
Ratio: 4.077 Gaps: 0
Percent Similarity: 76.471 Percent Identity: 58.824
alignment_block:
US-09-215-435-179_COPY_1_23 x US-09-969-730-67/rev ..
Align seg 1/1 to reverse of: US-09-969-730-67 from: 1 to: 1323
6 TrpLeuLeuLeuPheLeuPhePheLeuPhePheLeuPheLeuThrar 22
||| |||:::||||||| |||||:::||||| ::: ||
474 TGGGCCCTTTCTCTCTCTCAGTTTCTTCTACCTCTTCAGGTCTCTCAG 425
22 g 22
424 G 424

alignment_scores:

us-09-215-435-179_copy_1_23.rnpn

Thu Nov 15 10:52:04 2001

OM of: US-09-215-435-179_COPY_1_23 to: EST: * out_format : pfs
Date: Nov 14, 2001 10:30 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.mcdl -DEV=xlp
-O=/cpn1_1/USPto.spool/US09215435/runat_14112001_141104_2118/app_query.fasta_1.1519
-DB=EST -QFW=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFWT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435_ECGLN1_6733 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179_COPY_1_23
Query length: 23
Database: EST: *
Database sequences: 10228115
Database length: 431459454
Search time (sec): 5896.450000

score_list:	Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_gss24:AZ337349	-	67.00	154.32	17.77	525	I AZ337349 IM0068D04F Mouse 10kb
gb_est87:BA215978	-	69.00	131.85	24.38	406	I BF415978 UI-R-CAO-bxp-e-12-0-UI
gb_gss31:AZ711429	+	67.00	131.49	25.53	425	I AZ711429 RPCI-24-71024-TJ RPCI-1
gb_gss10:AZ0346731	-	67.00	148.84	35.88	595	I AQ346731 RPCI11-12517-TV RPCI-1
gb_gss20:AZ034197	+	66.00	137.04	12.52	158	I AZ034197 RPCI-23-377D15-TV RPCI
gb_gss20:AZ010385	+	66.00	149.51	32.92	411	I AZ010385 RPCI-23-365G17-TV RPCI
gb_gss28:AZ556218	+	66.00	149.30	33.81	422	I AZ556218 RPCI-23-177P7-TV RPCI
gb_est30:AU039151	-	65.00	153.30	20.23	191	I AU039151 AU039151 Dictyostelium
gb_est21:AU490726	-	65.00	149.14	34.52	324	I AU490726 EST241434 tomato shoot
gb_est21:AU490728	-	65.00	148.39	37.97	356	I AU490728 EST241436 tomato shoot
gb_est97:BG129260	-	65.00	147.22	44.01	412	I BG129260 EST474906 tomato shoot
gb_est97:BG130556	-	65.00	147.22	44.12	413	I BG130556 EST474902 tomato shoot
gb_est48:AW560372	+	65.00	145.26	56.78	530	I AW560372 EST315420 DSIR Medicac
gb_gss13:AZ0554809	-	65.00	145.20	57.21	534	I AQ554809 RPCI-11-382G13-TJ RPCI
gb_est103:C91466	-	65.00	144.71	60.89	568	I C91466 C91466 Dictyostelium dis
gb_est80:BE910551	-	65.00	143.76	68.81	641	I BE910551 601501101F1 NIH_MGC_70
gb_est93:BF867577	-	64.50	139.07	125.59	1008	I BF867577 963092F09.y1 C. reinh
gb_est71:BE158828	-	64.00	144.83	60.00	421	I BE158828 IL2-HT0397-151299-026-
gb_est34:AZ904272	+	64.00	144.48	62.74	440	I AZ904272 RPCI-24-221F13-TJ RPCI
gb_est22:AU1616040	+	64.00	144.04	66.34	465	I AU1616040 mg6808.y1 Soares.thym
gb_gss18:AZ0936195	+	64.00	143.09	75.00	525	I AQ936195 RPCI-11-Seg1-4-608P20-
gb_gss25:AZ373346	+	64.00	142.44	81.51	570	I AZ373346 IM012505R Mouse 10kb
gb_gss26:AZ409528	+	64.00	142.43	81.65	571	I AZ409528 IM0181L11F Mouse 10kb
gb_gss32:AZ790560	+	64.00	140.62	102.93	718	I AZ790560 2M0039E12F Mouse 10kb
gb_est76:BE615675	+	63.00	140.23	108.15	754	I BE615675 601279360F1 NIH_MGC_39
gb_est108:AU560483	+	63.00	146.08	51.08	270	I AU560483 V131F10.r1 Stratagene
gb_est108:AU560483	+	63.00	146.08	51.08	270	I T07865 EST05755 Fetal brain, St
gb_est20:AU1437332	+	63.00	143.35	72.54	382	I AU1437332 fb30504.x1 zebrafish w
gb_gss29:AZ619755	+	63.00	140.92	99.08	520	I AZ619755 IM0452108F Mouse 10kb
gb_gss27:AZ477895	+	63.00	139.37	120.87	633	I AZ477895 IM0297C04R Mouse 10kb
gb_gss1:AA550316	+	63.00	137.99	144.25	754	I AA550316 1457m3 gmbpFHB3.1, G.
gb_est103:C23639	-	63.00	136.31	178.91	933	I C23639 C23639 Dictyostelium dis
gb_gss16:AU0750720	-	63.00	135.90	188.60	983	I AU0750720 HS-5574_B2_G03-T7A RPC
gb_gss16:AU0750720	-	63.00	135.90	188.60	983	I AQ780342 HS_3184_A2_D07-T7C CI
em_estp19:BF262994	-	63.00	133.44	258.36	1342	I BF262994 HV_CEA0005I08F Horden
gb_est37:AV356560	-	62.50	149.39	33.44	154	I AV356560 AV356560 RIKEN full-le
gb_est28:AL378550	-	62.50	149.44	105.31	479	I AL378550 MCB38H07F1 MBBB Medic
gb_est89:BF520412	-	62.50	138.93	127.78	580	I BF520412 EST457882 DSIL Medicac
gb_gss27:AZ474525	-	62.50	138.61	132.46	601	I AZ474525 IM0292M02F Mouse 10kb
gb_est73:BE325277	-	62.50	138.01	143.82	652	I BE325277 NF119D06STIF1057 Devel

gb_gss29:AZ611033 - 62.50 137.26 158.33 717 I AZ611033 IM0436L02R Mouse 10
gb_est15:AU066195 + 62.00 147.82 40.90 163 I AU066195 TENU2586 T. cruzi e
gb_gss21:AZ100984 - 62.00 144.80 60.22 239 I AZ100984 RPCI-23-480N14-TV R
gb_est100:BG426365 - 62.00 138.86 129.06 508 I BG426365 602493552F1 NIH_MGC
seq_name: gb_gss24:AZ337349
seq_documentation_block:
LOCUS AZ337349 525 bp DNA 29-SEP-2000
DEFINITION IM0068D04F Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0068D04 F, DNA sequence.
ACCESSION AZ337349
VERSION AZ337349.1 GI:10407556
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 525)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dbunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0068 row: D column: 04
Seq primer: CGTTGTAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 525.
FEATURES
Location/Qualifiers
1..525
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0068D04"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 126 a 146 c 157 g 95 t 1 others
ORIGIN
alignment_scores: 69.00 Length: 22
Quality:

us-09-215-435-179_copy_1_23.rst

Thu Nov 15 10:52:04 2001

```

TAG_LIB=UI-R-CAO
TAG_TISSUE=corpus-striatum
TAG_SEQ=CTAGG*

BASE COUNT      116 a 102 c 74 g 114 t
ORIGIN

alignment_scores:
  Quality: 67.00      Length: 27
  Ratio: 3.045        Gaps: 1
  Percent Similarity: 81.481  Percent Identity: 55.556

alignment_block:
US-09-215-435-179_COPY_1_23 x BF415978/rev ..
Align seg 1/1 to reverse of: BF415978 from: 1 to: 406

1 MetMetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLe 17
  ::::::::::::::::::::::::::::::::::::::::::::::
161 CTCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 112

17 uPheLeu.....LeuThrArgGly 23
  ::::::::::::::::::::::::::::::::::::::::::::::
111 CCTCTCTCATTTTCGGAGGTGCACAGCGGA 81

seq_name: gb_gss31:AZ714249

seq_documentation_block: 425 bp DNA GSS 24-JAN-2001
LOCUS AZ714249
DEFINITION RPCI-24-71J24.TJ RPCI-24 Mus musculus genomic clone RPCI-24-71J24,
DNA sequence.
ACCESSION AZ714249
VERSION AZ714249.1 GI:12449772
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 425)
AUTHORS Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaod@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 71 row: J Column: 24
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
  1..425
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="RPCI-24-71J24"
  /clone_lib="RPCI-24"
  /sex="Male"
  /cell_type="Spleen/Brain"
  /note="Vector: pTARBAC1"
  /site_1="BamHI; Site_2: BamHI;
  RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
  library was cloned in the pTARBAC1 cloning vector at the
  BamHI sites using MboI partially digested male C57BL/6J
  DNA."

FEATURES
  source
    1..406
    /organism="Rattus norvegicus"
    /strain="Sprague-Dawley"
    /db_xref="taxon:10116"
    /clone="UI-R-CAO-bkp-e-12-0-UI"
    /clone_lib="UI-R-CAO"
    /lab_host="DH10B (Life Technologies)"
    /note="Vector: pT73D-Pac (Pharmacia) with a modified.
    polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CAO
    library is a subcloned library derived from the following
    tissues: thalamus, cerebellum, hypothalamus, medulla, pons
    midbrain, cerebral cortex, corpus striatum, testis, and
    hippocampus. For a detailed description of the library
    from which this clone was derived, please visit our web
    site at ratseq.eng.ualowa.edu. The subtraction has been
    previously described in (Bonaldo, Lennon and Soares,
    Genome Research 6:791-806, 1996)

seq_documentation_block: 406 bp mRNA EST 28-NOV-2000
LOCUS BF415978
DEFINITION UI-R-CAO-bkp-e-12-0-UI.s1 UI-R-CAO Rattus norvegicus cDNA clone
UI-R-CAO-bkp-e-12-0-UI 3', mRNA sequence.
ACCESSION BF415978
VERSION BF415978.1 GI:11403967
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 406)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized-corpus-striatum library cDNA library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 3-129, >RMR1B#Unknown
130-166, >(GGA)n#Simple-repeat
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
  1..406
  /organism="Rattus norvegicus"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-CAO-bkp-e-12-0-UI"
  /clone_lib="UI-R-CAO"
  /lab_host="DH10B (Life Technologies)"
  /note="Vector: pT73D-Pac (Pharmacia) with a modified.
  polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CAO
  library is a subcloned library derived from the following
  tissues: thalamus, cerebellum, hypothalamus, medulla, pons
  midbrain, cerebral cortex, corpus striatum, testis, and
  hippocampus. For a detailed description of the library
  from which this clone was derived, please visit our web
  site at ratseq.eng.ualowa.edu. The subtraction has been
  previously described in (Bonaldo, Lennon and Soares,
  Genome Research 6:791-806, 1996)

```



```

seq_documentation_block:
LOCUS      AU039151      191 bp      mRNA      EST      29-MAR-1999
DEFINITION AU039151 Dictyostelium discoideum SS (H.Urushihiara) Dictyostelium
discoideum cDNA clone SSM450, mRNA sequence.
ACCESSION  AU039151
KEYWORDS   AU039151.1 GI:3985904
SOURCE     EST.
ORGANISM   Dictyostelium discoideum.
REFERENCE  1 (bases 1 to 191)
AUTHORS    Morio,T., Urushihiara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE      The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL    DNA Res. 5 (6), 335-340 (1998)
MEDLINE    99156227
COMMENT    Contact: Hideko Urushihiara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = "Dictyostelium discoideum cDNA project in Japan".
FEATURES   Location/Qualifiers
source     1..191
            /organism="Dictyostelium discoideum"
            /strain="AX4"
            /db_xref="taxon:44689"
            /clone="SSM450"
            /clone_lib="Dictyostelium discoideum SS (H.Urushihiara)"
            /dev_stage="slug"
BASE COUNT 130 a      8 c      8 g      45 t
ORIGIN
alignment_scores:
Quality: 65.00      Length: 18
Ratio: 4.062      Caps: 0
Percent Similarity: 88.889      Percent Identity: 72.222
alignment_block:
US-09-215-435-179_COPY_1_23 x AU039151/rev ..
Align seg 1/1 to reverse of: AU039151 from: 1 to: 191
2 MetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLeuPh 18
:::|||||
97 TTACTATATTACTGGTTATTATTATTATTATTATTATTATTTT 48
18 eLeu 19
||||
47 TTTA 44
seq_name: gb_est21:AI490726
seq_documentation_block:
LOCUS      AI490726      324 bp      mRNA      EST      29-JUN-1999
DEFINITION EST241434 tomato shoot, Cornell Lycopersicon esculentum cDNA clone
cLEB3L10, mRNA sequence.
ACCESSION  AI490726
VERSION    AI490726.1 GI:4386036
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE  1 (bases 1 to 324)
AUTHORS    van der Hoeven,R.S., Matern,A.L., Vision,T., Holt,I.E., Liang,F.,
Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE      Generation of ESTs from tomato shoot meristem
JOURNAL    Unpublished (1999)
COMMENT    Other_ESTs: EST241434
Contact: David Frisch

```

```

Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
Generation of ESTs from tomato shoot meristem
Unpublished (1999)
Other_ESTs: EST241436
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
FEATURES   Location/Qualifiers
source     1..324
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEB3L10"
            /clone_lib="tomato shoot, Cornell"
            /tissue_type="shoot meristem"
            /dev_stage="8 week old plants"
            /lab_host="XLOLR"
            /note="vector: pBK_CMV; Site_1: EcoRI; Site_2: XhoI; cLEB
- Tomato Shoot Meristem EST Library. Oligo-dT primed cDNA
library made from tomato vegetative shoots including
meristems and small expanding leaves."
BASE COUNT 103 a      43 c      105 g      73 t
ORIGIN
alignment_scores:
Quality: 65.00      Length: 19
Ratio: 3.824      Caps: 0
Percent Similarity: 89.474      Percent Identity: 57.895
alignment_block:
US-09-215-435-179_COPY_1_23 x AI490726/rev ..
Align seg 1/1 to reverse of: AI490726 from: 1 to: 324
2 MetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLeuPh 18
:::|||||
176 CTTTTCAGCAGGTGGTGTGATCCTCGTCTCCCTCTCTCTCTCT 127
18 eLeuLeu 20
|||||
126 CCTCATC 120
seq_name: gb_est21:AI490728
seq_documentation_block:
LOCUS      AI490728      356 bp      mRNA      EST      29-JUN-1999
DEFINITION EST241436 tomato shoot, Cornell Lycopersicon esculentum cDNA clone
cLEB3L10, mRNA sequence.
ACCESSION  AI490728
VERSION    AI490728.1 GI:4386038
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE  1 (bases 1 to 356)
AUTHORS    van der Hoeven,R.S., Matern,A.L., Vision,T., Holt,I.E., Liang,F.,
Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE      Generation of ESTs from tomato shoot meristem
JOURNAL    Unpublished (1999)
COMMENT    Other_ESTs: EST241434
Contact: David Frisch

```

```

/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT      128 a      53 c      118 g      113 t
ORIGIN

alignment_scores:
    Quality:      65.00      Length:      19
    Ratio:        3.824      Gaps:      0
Percent Similarity: 89.474      Percent identity: 57.895

alignment_block:
US-09-215-435-179_COPY1_1_23 x BG129260/rev ..
Align seg 1/1 to reverse of: BG129260 from: 1 to: 412

2 MetLeuProGlnTrpLeuLeuLeuPheLeuLeuPhePheLeuPh 18
      :::::  ::::::::::::::::::::  ::::::::::::::
182 CTTTTCAGCAGGTGTTGATGCTCGTCTTCTCTCGTCTTCTTCTTCTT 133

18 eLeuLeu 20
      :::::
132 CCTCATC 126

seq_name: gb_est97:BG130556

seq_documentation_block:
LOCUS      BG130556      413 bp      mRNA      EST      31-JAN-2001
DEFINITION      EST476202 tomato shoot/meristem Lycopersicon esculentum cDNA clone
               CTOF30P11 5' sequence, mRNA sequence.
ACCESSION      BG130556
VERSION        BG130556.1 GI:12630744
KEYWORDS
SOURCE        tomato.
ORGANISM      Lycopersicon esculentum
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
               Lycopersicon.
REFERENCE      1 (bases 1 to 413)
AUTHORS      van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
               Hansen,C., Ronning,C. and Tanksley,S.
TITLE        Generation of ESTs from tomato shoot/meristem tissue
JOURNAL      Unpublished (2001)
COMMENT      Contact: CUGI
               Clemson University Genomics Institute
               Clemson University
               100 Jordan Hall, Clemson, SC 29634, USA
               Email: http://www.genome.clemson.edu/orders/index.html.
               Location/Qualifiers
FEATURES             1..413
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="CTOF30P11"
                     /tissue_type="shoot/meristem"
                     /dev_stage="developing shoots from 4-6wks old plants"
                     /lab_host="SOLR"
                     /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Small expanding leaves from the growing tip were
                     taken from greenhouse plants (4-6wks old TA496). Tissue
                     was immediately frozen in liquid nitrogen."
BASE COUNT      128 a      53 c      118 g      114 t
ORIGIN

```

```

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
Location/Qualifiers
1. .356
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="cLEB3L10"
/clone.lib="tomato shoot, Cornell"
/tissue_type="shoot meristem"
/dev_stage="8 week old plants"
/lab_host="XLOR"
/note="vector: pBK CMV; Site_1: EcoRI; Site_2: XhoI; cLEB
-tomato Shoot Meristem EST library. Oligo-dr primed CDNA
library made from tomato vegetative shoots including
meristems and small expanding leaves."
91 a lll c 46 g 108 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 65.00 Length: 19
Ratio: 3.824 Gaps: 0
Percent Similarity: 89.474 Percent Identity: 57.895

alignment_block:
US-09-215-435-179_COPY_1_23 x AI490728 ..
Align seg 1/1 to: AI490728 from: 1 to: 356

2 MetLeuProGIntTPLeuLeuLeuPheLeuPhePhePhLeuPh 18
+++++:::|||||:::|||||:::|||||:::|||||:::|||||
181 CTTTACAGGTGGTGATCGCTTCGCTGCCTCTCTCTCTCTT 230

18 eLeuLeu 20
||||:::
231 CCTCATC 237

seq_name: gb_est97:BG129260

seq_documentation_block: 412 bp mRNA EST 31-JAN-2001
LOCUS BG129260
DEFINITION EST474906 tomato shoot/meristem Lycopersicon esculentum cDNA clone
ctOF23f6 5' sequence, mRNA sequence.
ACCESSION BG129260
VERSION BG129260.1 GI:12629448
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
1 (bases 1 to 412)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Konning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1. .412
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="ctOF23f6"
FEATURES
source
```


alignment_scores:

Quality: 65.00 Length: 19
Ratio: 3.824 Gaps: 0
Percent Similarity: 89.474 Percent Identity: 57.895

alignment_block:

US-09-215-435-179_COPY_1_23 x BG130556/rev ..

Align seg 1/1 to reverse of: BG130556 from: 1 to: 413

2 MetLeuProGInTrpLeuLeuLeuPheLeuPhePheLeuPh 18

182 CTTTTCAGAGGTGGTGCCTCGCTTCCTCCCTGCTCTCTT 133

18 eleuLeu 20

||||:

132 CCTCATC 126

seq_name: gb_est48:AW560372

seq_documentation_block:

LOCUS AW560372 530 bp mRNA EST 07-SEP-2000
DEFINITION EST315420 DSIR Medicago truncatula cDNA clone pDSIR-26N4, mRNA

sequence.

ACCESSION AW560372

VERSION AW560372.1 GI:7205798

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.

REFERENCE 1 (bases 1 to 530)

AUTHORS Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.

TITLE ESTs from roots of Medicago truncatula after inoculation with

Phytophthora medicaginis

Unpublished (1999)

JOURNAL Contact: Carroll P. Vance

COMMENT Department of Agronomy and Plant Genetics

University of Minnesota

411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA

Tel: 612 625 5715

Fax: 651-649-5058

Email: vance004@maroon.tc.umn.edu

Minnesota EST name:M251727e ; TIGR sequence name:MTBAX74TK ; More

information, including clone ordering, is available at .

'http://chryslie.tamu.edu/medicago'

Seq primer: SKmod (CTA gAA CTA gtg gAT CC).

FEATURES

source

1..530

/organism="Medicago truncatula"

/cultivar="genotype A17"

/db_xref="taxon:3880"

/clone="pDSIR-26N4"

/clone_lib="DSIR"

/tissue_type="roots infected with Phytophthora

medicaginis"

/dev_stage="roots harvested at 10 days post inoculation

with Phytophthora medicaginis"

/lab_host="E. coli strain XLOLR"

/note="vector: pBluescript SK-; Site.1: EcoRI; Site.2:

XhoI; cDNA was prepared from polyA+ enriched RNA from

roots harvested at 10 days post inoculation with

Phytophthora medicaginis. The cDNA was directionally

ligated into the Uni-ZAP XR vector from Stratagene and

packaged using Gigapack III Gold packaging extracts.

Plasmids containing cDNA inserts were excised from the

recombinant lambda-ZAP phage using Ex-Assist helper phage

origin."

BASE COUNT 86 a 147 c 39 g 258 t

ORIGIN

alignment_scores:
Quality: 65.00 Length: 20
Ratio: 3.421 Gaps: 0
Percent Similarity: 95.000 Percent Identity: 65.000

alignment_block:

US-09-215-435-179_COPY_1_23 x AW560372

Align seg 1/1 to: AW560372 from: 1 to: 530

1 MetMetLeuProGInTrpLeuLeuLeuPheLeuPhePheLeu 17

||||:

236 CTTATCCTTTCAGATTCTCTCTCTCTCTCTCTCTCTTTCT 285

17 uPheLeuLeu 20

||||:

286 TCTCTCTCTC 295

seq_name: gb_gss13:AO554809

seq_documentation_block:

LOCUS AO554809 534 bp DNA GSS 28-MAY-1999
DEFINITION RPCI-11-382G13.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-382G13
, DNA sequence.

ACCESSION AO554809

VERSION AO554809.1 GI:4913986

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 534)

AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

JOURNAL Other_GSSs: RPCI-11-382G13.TV

COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6

Class: BAC ends.

FEATURES

source

1..534

/organism="Homo sapiens"

/db_xref="GDB:764640"

/db_xref="taxon:9606"

/clones="RPCI-11-382G13"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT 216 a 64 c 141 g 113 t

ORIGIN

alignment_scores:
Quality: 65.00 Length: 19


```

alignment_scores:
  Quality: 343.00      Length: 340
  Ratio: 3.811        Gaps: 2
  Percent Similarity: 26.471  Percent Identity: 25.882

alignment_block:
US-09-215-435-179_copy_24_121 x AL157823 ..

Align seg 1/1 to: AL157823 from: 1 to: 127384

      8 AsnLeuLeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluTh 24
      87881 TCCCTGCAGGAGCTCAAGAGCTCTGTCATCGGAACCAAGGACTCGGAGAC 87930
      24 rGlyCysCysGlnArGAlaProAspAsnCysGluSerHisCysAlaGluL 41
      87931 TGGCTGCTGCCAACGCTGCTCCAGACAATGGCAGTCGCAGTCGCGGGAGA 87980
      41 ySgLySerGluGlySerLeuCysGlnThrGlnVal..Phe..... 53
      87981 AGGGGTCCGAGGCGACTCTGCTCAACGCGAGCTGGGTATCGCGGCCCGG 88030
      53 ..... 53
      88031 GGGGAGCCAGAGGGGATCCAGGGGAAGTGGGAGCCAGCGCGGCCAGAT 88080
      53 ..... 53
      88081 TCCTGGGGAGGAAGAAGAGGTAGGTGGCGGAATGCCCTGGAGCCTGGA 88130
      53 ..... 53
      88131 ATTCCCCCGTGGGACTGCACCTCCAAGCCTTCTGGGGAGCCTGTGGACG 88180
      53 ..... 53
      88181 GAACCAATATGGTAGCCACCCTAGGCTTTTGGATTCCTGGATGGGGCG 88230
      53 ..... 53
      88231 GGAAGGCGATTTTCACTTTCTAATGTGGTTCCACATTTCTCCAGAGTCT 88280
      53 ..... 53
      88281 TCTCAGCAAAATCCGAAAGTAGTACAGTAATGGGCACAGCTAGGTCCAG 88330
      53 ..... 53
      88331 CCTCTGCCCCAGAGCCGCTCTTCATCCAGCTCAGGAGTCTGCTGTGGC 88380
      53 ..... 53
      88381 CCCGGGTGTTCTGTTGAGGATGGGCAGGACGCCCTCTGAGTGGGTG 88430
      53 ..... 53
      88431 TGTAGGGGAGGAGTGTGATTTGCCCATGGAAGGATTTAAAGACAAGACAT 88480
      53 ..... 53
      88481 AGAGGTCTTCGATGTTGTGTCTCCATGGACTTAGGACAGGACATGGGTCA 88530
      53 ..... 53
      88531 GTGACAGCAGGGTTACAAAGGTTGAGGCTGGAGTGGGAGGACACTGAACAATCC 88580
      53 ..... 53
      88581 AGGAGCCAGCCTGGAGAGGACGCCAGAGGTGCCACTCTGACGACCTGGC 88630
      53 ..... 53
      88631 AAGCTATCATTTATCCTGGATCCTGGTGTGGGCAAGCACATGGGCCCT 88680

```

```

53 ..... 53
88681 GGAGCTGGGCTGGGCTTACGGGCTCAGGGAAAGAGCTGCCCCCAAG 88730
54 .....PheGlyGlnTyrArgA 59
      88731 GTCGAGGTCMAAGTCTCTTCCAGGTGTTCTTGGCCAATATAGAG 88780
59 laCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 75
88781 CGTGTCCCTCGCTGGGAACCTGACTTGTATATATTCAAGAAGATGAGAAA 88830
76 TrpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLe 92
88831 TGGCTTAGCATCGCTATGCGCGTGTGTCAAGAAATTTGGAAGGCAGAACTT 88880
92 uAlaLysLysMetPhePhe 98
88881 GGCTAAGAAAATGTTCTTC 88899
seq_name: gb_htg22:AL512371
seq_documentation_block:
AL512371 159482 bp DNA HTG 21-MAR-2001
LOCUS Homo sapiens chromosome 6 clone RP11-238F24, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION AL512371
VERSION AL512371.4 GI:13274864
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 159482)
Clark, G.
Direct Submission
Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequest@sanger.ac.uk
On Mar 12, 2001 this sequence version replaced gi:12964529.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA238F24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 154970 bases at least Q40
Consensus quality: 156246 bases at least Q30
Consensus quality: 156738 bases at least Q20
Insert size: 159282; sum-of-contigs
Insert size: 165487; 7.5% error; agarose-fp
Quality coverage: 6.35x in Q20 bases; sum-of-contigs Quality
coverage: 6.11x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved. 3066: contig of 3066 bp in length
* 1
* 3067 3166: gap of 100 bp
* 3167 155002: contig of 151836 bp in length
* 155003 155102: gap of 100 bp
* 155103 159482: contig of 4380 bp in length.

```

FEATURES
source Location/Qualifiers
1. .159482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-238F24"
1. .3066
/note="assembly_fragment:02384
fragment_chain:1"
misc_feature 3167..155002
/note="assembly_fragment:02553
fragment_chain:1"
155103..159482
/note="assembly_fragment:02886"
BASE COUNT 42530 a 37388 c 36532 g 40831 t 201 others
ORIGIN

alignment_scores:
Quality: 343.00 Length: 340
Ratio: 3.811 Gaps: 2
Percent Similarity: 26.471 Percent Identity: 25.882

alignment_block:
US-09-215-435-179_COPY_24_121 x AL512371 ..
Align seg 1/1 to: AL512371 from: 1 to: 159482

8 AsnLeuLeuGluLeuLysGluSerCysLeuArgAsnGlnAspCysGluTh 24
:::|||||
101423 TCCCTGCAGGAGCTCAAGAGTCTTGCATCCGGAACCAAGGACTCGGAGAC 101472

24 rGlyCysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluL 41
|||||
101473 TGGCTGCTGCCAACGTGCTCCAGACAAATGCCAGTCGCACCTCGCGGAGA 101522

41 ysGlySerGluGlySerLeuCysGlnThrGlnVal..Phe..... 53
|||||
101523 AGGGGTCCGAGGCGACTGTGTCTCAACGCGAGGTGGGTATCCCGCCCGG 101572

53 53

101573 GGGGAGCCAGAGGGGATCCAGGGGAAGTGGAGCCAGGAGGGCCAGAT 101622

53 53

101623 TCCTGGGGAGGAAGAAGGGTAGTGGGCGGAAATGCCCTGGAGCCTGGA 101672

53 53

101673 ATTCCCCCTGGGACTGCACCTCCAAAGCCTTCTGGGGGAGCCTGTGGAGC 101722

53 53

101723 GAACCAATATGTAGCCACCCTAGGGTTTGGAAATCCTGGATTGGGGCG 101772

53 53

101773 GGAAGCGGATTTTCACTTTCTAATGTGGGTTTCACATTCCTCAGGAGTCT 101822

53 53

101823 TCTCAGCAAAATTCGNAAGTAGTACAGTAATGGGCACAGCTAGGTCCAG 101872

53 53

101873 CCTCTGCCCCAGAGCCCCGTCTTTCATCCAGCTCAGGAGTTCTGCTGTGC 101922

53 53

101923 CCCGGGTGTTCTCTGTGGAGGATGGGCAGAGGAGCCTCTGAGCTGGGTG 101972

53 53
101973 TCTAGGGGAGGAGTGTGATTTGCCCATGAAGGATTTAAAGACAGACAT 102022

53 53
102023 AGAGTCTCCTGCGATGGTGTGTCCATGGACTTAGGACAGCAGATGGGTCA 102072

53 53
102073 GTGACAGGCGAGGTTACAGGTTCTGGAGCTGGAGGACACTGAACAATCC 102122

53 53
102123 AGGAGCCAGCCTTGAGAGAGGAGCCAGAGGTGCCACTCTGAGCAGCTGGC 102172

53 53
102173 AAGCTATCATTTATCCTTGGATCCTGGCTGTGGCAAGCACATGGGCCCT 102222

53 53
102223 GGAGCTGGCTGGCTGGCTTAGGGGCTCAGGGAAGACTGCCGCCAAG 102272

54 54
102273 GTCGAGGTCAAAGTCTGTCCTTTCCAGGCTGTTCTTTGCCCAATATAGAG 102322

59 laCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 75
|||||
102323 CGTGTCCCTGCTGGGGAACCTGACTTGTATATATTCAAAGAATGAGAAA 102372

76 TtpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLe 92
|||||
102373 TGGCTAGCATCGCCTATGCGCGCTGTGCAGAAATTTGGAAGGCAGGAATT 102422

92 uAlaLysLysMetPhePhe 98
|||||
102423 GCCTAAGAAATGTTCTTC 102441

seq_name: gb_htg23:AL590400

seq_documentation_block:
LOCUS AL590400 168884 bp DNA HTG 08-APR-2001
DEFINITION Homo sapiens chromosome 6 clone RP11-569I12, *** SEQUENCING IN
PROGRESS ***, 14 unordered pieces.
ACCESSION AL590400
VERSION AL590400.2 GI:13568356
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Sims,S.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:13559131.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA569I12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 164511 bases at least Q40
Consensus quality: 165952 bases at least Q30

Consensus quality: 166791 bases at least Q20
Insert size: 167584; sum-of-contigs
Insert size: 185239; 2.7% error; agarose-fp
Quality coverage: 6.19x in Q20 bases; sum-of-contigs Quality
Coverage: 5.74x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 2920: contig of 2920 bp in length
* 2921 3020: gap of 100 bp
* 3021 32168: contig of 29148 bp in length
* 32169 32268: gap of 100 bp
* 32269 36080: contig of 3812 bp in length
* 36081 36180: gap of 100 bp
* 36181 45704: contig of 9524 bp in length
* 45705 45804: gap of 100 bp
* 45805 59680: contig of 13876 bp in length
* 59681 59780: gap of 100 bp
* 59781 61890: contig of 2110 bp in length
* 61891 61990: gap of 100 bp
* 61991 72185: contig of 10195 bp in length
* 72186 72285: gap of 100 bp
* 72286 76243: contig of 3958 bp in length
* 76244 76343: gap of 100 bp
* 76344 81840: contig of 5497 bp in length
* 81841 81940: gap of 100 bp
* 81941 86241: contig of 4301 bp in length
* 86242 86341: gap of 100 bp
* 86342 129124: contig of 42783 bp in length
* 129125 129224: gap of 100 bp
* 129225 131985: contig of 2761 bp in length
* 131986 132085: gap of 100 bp
* 132086 151439: contig of 19354 bp in length
* 151440 151539: gap of 100 bp
* 151540 168884: contig of 17345 bp in length.
*
* Location/Qualifiers
* 1..16884 /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="6"
* /clone_lib="RP11-569I12"
* /clone_lib="RP11-11.2"
* 1..2920 /note="assembly_fragment:00514"
* fragment_chain:1
* clone_end:SP6
* vector_side:left
* 3021..32168 /note="assembly_fragment:01440"
* fragment_chain:1
* 32269..36080 /note="assembly_fragment:00777"
* fragment_chain:1
* 36181..45704 /note="assembly_fragment:00971"
* fragment_chain:1
* 45805..59680 /note="assembly_fragment:00349"
* fragment_chain:1
* 59781..61890 /note="assembly_fragment:02150"
* fragment_chain:2
* 61991..72185 /note="assembly_fragment:02579"
* fragment_chain:2
* 72286..76243 /note="assembly_fragment:00659"

misc_feature
76344..81840
/note="assembly_fragment:00425"
fragment_chain:2
81941..86241
/note="assembly_fragment:02764"
fragment_chain:2
86342..129124
/note="assembly_fragment:00846"
fragment_chain:2
129225..131985
/note="assembly_fragment:01963"
132086..151439
/note="assembly_fragment:02120"
fragment_chain:3
151540..168884
/note="assembly_fragment:01758"
fragment_chain:3
clone_end:T7
vector_side:right
BASE COUNT 46011 a 38577 c 39446 g 43544 t 1306 others
ORIGIN

alignment_scores:
Quality: 343.00 Length: 340
Ratio: 3.811 Gaps: 2
Percent Similarity: 26.471 Percent Identity: 25.882

alignment_block:
US-09-215-435-179_copy_24_121 x AL590400 ..
Align seg 1/1 to: AL590400 from: 1 to: 168884

8 AsnLeuLeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThr 24
|||||
114919 TCCCTGCGAGGAGCTCAAGGAGTCTTGCATCGGAACCCAGGACTCGGAGAC 114968
|||||
24 rGlyCysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluL 41
|||||
114969 TGGCTGCTGCCAACGTCGTCCAGACAAATTGGCAGTTCGCAGTCGCCGGAGA 115018
|||||
41 ysGlySerGluGlySerLeuCysGlnThrGlnVal..Phe..... 53
|||||
115019 AGGGGTCCGAGGGCAGTCTGTGTCAAACGACAGGTGGGTATCCGCCGCCGG 115068
|||||
53 53

115069 GGGGAGCCAGAGGGGATCCAGGGGAAGTGGGAGCCAGGGAGGCCACAT 115118
53 53

115119 TCCTGGGGAGGAAGAAGGGTAGGTGGGGCGGAATGCCCTGGAGCCTGGA 115168
53 53

115169 ATTCCCCCGTGGGAGTGCACCTCCAAGCCTTCTGGGGGAGCCTGTGGAG 115218
53 53

115219 GAACCAATATGTAGCCACCCCTAGGGTTTGGATCTCTGGATTGGGGCG 115268
53 53

115269 GGAAGGGGATGTTTCACTTTTCTAATGTGGGTTCAATTTCTCCAGGAGTCT 115318
53 53

115319 TCTCAGCAAAATCCGAAAGTAGTACAGTAATAATGGGCACAGCTAGGTCCAG 115368
53 53

115369 CCTCTGCCCCAGAGCCCCGTTTTCATCCAGCTCAGGAGTTCTGCTGTGGC 115418

```
53 ..... 53
115419 CCCGGGTGTTCTGTTGGAGGATGGCAGAGCAGCCCTCTGAGCTGGGTG 115468
53 ..... 53
115469 TGTAGGGAGGGAGTTGATTGTCCTCATGAAGGATTTAAAGACAAGACAT 115518
53 ..... 53
115519 AGAGTCTCGCATGCTGTGTCATGGACTTAGGACAGCAGATGGGTCA 115568
53 ..... 53
115569 GTGACAGCAGGGTTACRAGGTCTGGAGCTGGGAGGACACTGAACAATCC 115618
53 ..... 53
115619 AGGAGCCAGCTGGAGAGGACGACAGAGTGCCACTCTGAGCAGCTGGC 115668
53 ..... 53
115669 AAGCTATCATTTATCCGTGATPCTGTGGTGTGGCAAGCACATGGGCCCT 115718
53 ..... 53
115719 GGAGCTGGGCTGGGCTGGCTTAGGGCTCAGGGAAGAGTCCGCCCAAG 115768
54 ..... PheGlyGlnTyrArgA 59
115769 GTGAGGTCAAAGTCTGCTTTCCAGGTGTTCTTGGCCAATATAGAG 115818
59 laCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 75
115819 COTGTGCTGCTGCGGAACCTGACTTGTATATATCAAGAATGAGAA 115868
76 TrpLeuSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLe 92
115869 TGGCTTAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115918
92 uAlaLysLysMetPhePhe 98
115919 GGCTAAGAAATGTTCTTC 115937

seq_name: gb_om:CCCOLIP

seq_documentation_block: 452 bp mRNA MAM 31-MAR-1995
LOCUS CCCOLIP
DEFINITION Canine mRNA for colipase.
ACCESSION X53564
VERSION X53564.1 GI:841
KEYWORDS colipase.
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Fukuoka,S.I. and Scheele,G.A.
Direct Submission
Submitted (19-JUN-1990) Fukuoka S.-I., Scheele G.A., The Charles A.
Dana Research Institute, Harvard Medical School, Cell and Molecular
Biology, Dana 533, Beth Israel Hospital, 330 Brookline Ave.,
Boston, MA 02215, USA
2 (bases 1 to 452)
Fukuoka,S., Taniguchi,Y., Kitagawa,Y. and Scheele,G.
Full length cDNA sequence encoding canine pancreatic colipase
Nucleic Acids Res. 18 (18), 5549 (1990)
MEDLINE 91016846
FEATURES
    Location/Qualifiers
        1..452
            /organism="Canis familiaris"
            /db_xref="taxon:9615"
```

```

mRNA
/map="6pter-p21.1"
<1. .523

```

TITLE Human nucleic acid sequences from

nucleic acid sequences from normal pancreatic islets

JOURNAL Patent: WO 9954446-A 16 28-OCT-1999;
BERNDT ARMIN (DE); SPECHT THOMAS (DE); HINZMANN
BERNDT (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source Location/Qualifiers
1. .656
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 165 a 200 c 162 g 129 t
ORIGIN
alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184
alignment_block:
US-09-215-435-179_COPY_24_121 x AX014220 ..
Align seg 1/1 to: AX014220 from: 1 to: 656
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||:||||: |||:||||: |||:||||: |||
170 CTGGAGAACGGTGCCTCTGCATGAATAGTCCCAAGTGTAGAGCAATTG 219
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 43
|||||:||||: |||:||||: |||:||||: |||:||||: |||
220 CTGCAGCATTCAGTGGCTGGCTGGCGCGCGCGACATCCATCGGCCA 269
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
|||||:||||: |||:||||: |||:||||: |||:||||: |||
270 GCGAGAACAGCGAGTCTGTCTCAAGACGCTCTATGGGATTACTACAG 319
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
|||||:||||: |||:||||: |||:||||: |||:||||: |||
320 TGTCCCTGTGAGCTGGCGCTGACCTGT.....GAGGGAGACAAGAC 360
76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 88
|||:||||: |||:||||: |||:||||: |||:||||: |||
361 CATCTGGGCTCATCACCAACCACTTGGCATCTGCCATGAGCTG 410
88 lyArgGlnLys 91
|||||:||||: |||:||||: |||:||||: |||:||||: |||
411 GAGCGTCCAAG 421
seq_name: gb_rol:MCCOLIPAS
seq_documentation_block:
LOCUS MCCOLIPAS 490 bp mRNA ROD 21-NOV-1996
DEFINITION M. coypus mRNA for colipase.
ACCESSION X82998
VERSION X82998.1 GI:599866
KEYWORDS colipase.
SOURCE
ORGANISM Myocastor coypus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystricognathi; Myocastoridae;
Myocastor.
1 (bases 1 to 490)
Thirstrup,K., Carriere,F., Hjorth,S.A., Rasmussen,P.B.,
Nielsen,P.F., Ladefoged,C., Thim,L. and Boel,E.
Cloning and expression in insect cells of two pancreatic lipases
and a procollipase from Myocastor coypus
Eur. J. Biochem. 227 (1-2), 186-193 (1995)
95154288
MEDLINE
2 (bases 1 to 490)
Thirstrup,K.
AUTHORS
Direct Submission
TITLE
Submitted (28-NOV-1994) K. Thirstrup, Univ. Hospital, Lab. for
Molecular Endocrinology, Rigshospitalet 6321, Blegdamsvej 9,
DK-2100 Copenhagen, DENMARK
JOURNAL
FEATURES
Location/Qualifiers
seq_name: gb_om:ECCOLIPB
seq_documentation_block:
LOCUS ECCOLIPB 497 bp mRNA MAM 11-NOV-1994
DEFINITION E. caballus mRNA for colipase B.
ACCESSION X74344
VERSION X74344.1 GI:572678
KEYWORDS colipase.
SOURCE
ORGANISM Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE
1 (bases 1 to 497)
AUTHORS Kerfelec,B.
TITLE Direct Submission

JOURNAL Submitted (28-JUL-1993) B. Kerfelec, CNRS, CBM, 31 Chemin Joseph
REFERENCE Aiguier, 13402 Marseille, Cedex 9, FRANCE
AUTHORS 2 (bases 1 to 497)
TITLE Molecular cloning and expression of two horse pancreatic cDNA
encoding colipase A and B
JOURNAL Biochim. Biophys. Acta 1213 (3), 357-360 (1994)
MEDLINE 94325330
FEATURES Location/Qualifiers
source 1..497
/organism="Equus caballus"
/db_xref="taxon:9796"
/tissue_type="pancreas"
<1..327
/codon_start=1
/product="colipase B"
/protein_id="CAA52391.1"
/db_xref="GI:572679"
/db_xref="SWISS-PROT:P02705"
/translation="LALLLVALAAYAVDPGRVVIINLEAGEICMNSACKSECCCHRE
SISLARCARKASENSECSAWTLYGVYKPCRGTLQVDKTLVGSIMNTNFGICFD
AARSEE"
BASE COUNT 118 a 161 c 125 g 93 t
ORIGIN

alignment_scores:
Quality: 134.00 Length: 63
Ratio: 3.045 Gaps: 0
Percent Similarity: 69.841 Percent Identity: 39.683

alignment_block:
US-09-215-435-179_COPY_24_121 x ECCOLIPB ..
Align seg 1/1 to: ECCOLIPB from: 1 to: 497

```

10 LeuGluLeuLysClnSerCysIleArgAsnGlnAspCysGluThrGlyCys 26
||||| ||| |||:||||:||||:||||:||||:||||:
70 TTGGAGCGGGCGGAGATCTGCTGACACGCGCCAGTGCAAGAGCAATG 119
|||||:||||: ||: |||:||||: |||:||||: |||:||||:
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 43
|||||:||||: ||: |||:||||: |||:||||: |||:||||:
120 CTGCCACCGGAAAGCAGCGCTGAGCCTGGCCCGCTGCAGCAAGAGCCA 169
|||||:||||: ||: |||:||||: |||:||||: |||:||||:
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
|||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
170 GCAGAAACACGAGTGTCTGCTGGACACTCTACGGGGTTTACTACAAG 219
|||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLys 72
||||| |||:||||| |||:||||| |||:|||||
220 TGTCTCTGTGAGCGGCGCTGACCTGCCAGGTGGACAAG 258
||||| |||:||||| |||:||||| |||:|||||
```

seq_name: gb_om:AF148567

seq_documentation_block:
LOCUS AF148567 510 bp mRNA MAM 06-MAY-2000
DEFINITION Sus scrofa pancreatic colipase mRNA, complete cds.
ACCESSION AF148567
VERSION AF148567.1 GI:7711135
KEYWORDS
SOURCE plg.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 510)
AUTHORS Darnis,S. and Chaix,J.C.
TITLE Cloning, sequencing and functional expression of porcine pancreatic
colipase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 510)
AUTHORS Darnis,S. and Chaix,J.C.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) LBBN, CNRS-ESA6033, Av Escadrille Normandi

Niemen, Marseille 13397, France
Location/Qualifiers
1. 510
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="pancreas"
44..382
/note="cofactor; triglyceride lipase"
/codon_start=1
/product="colipase"
/protein_id="AAF67823.1"
/db_xref="GI:7711136"
/translation="MEKVLALLTLTVAYAVDPDPTIINLDELGELCLNSAOCKSKSC
QHDITLSLRCAALKARENSECFAFTLYGVVYKPCPERGLTCEGDKSILVGSIINTNFNG
ICHDVGRSSD"

BASE COUNT 117 a 160 c 121 g 112 t

ORIGIN

alignment_scores:
Quality: 133.00 Length: 85
Ratio: 2.463 Gaps: 2
Percent Similarity: 63.529 Percent Identity: 34.118

alignment_block:
US-09-215-435-179_COPY_24_121 x AFL148567 ..

Align seg 1/1 to: AFL148567 from: 1 to: 510

10 LeuGlLeuLySerCysLeuArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||||::: ::|||:::||||
125 CTGGATGAGCGGAGCTCTGCCTGAACAGTCGCCAGTGCAAGAGCAACTG 174
|||||::: |||||::: |||||::: |||||::: |||||:::
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLySgLy 43
|||||::: |||||::: |||||::: |||||::: |||||:::
175 CTGCCAGCATGACACAATCTGTAGCCGTCCCCGTGGCACCTCAAGGCCA 224
|||||::: |||||::: |||||::: |||||::: |||||:::
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArGAla 59
|||||::: |||||::: |||||::: |||||::: |||||:::
225 GAGAGAACACGAGTGTTCTGCCTTCACGCTCTATGGGGTTTACTACAAG 274
|||||::: |||||::: |||||::: |||||::: |||||:::
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLyAsnGluLyStr 76
|||||::: |||||::: |||||::: |||||::: |||||:::
275 TGTCCCTGTGAACGGGCGCTGACCTGT.....GAGGGGGACAAGAG 315
|||||::: |||||::: |||||::: |||||::: |||||:::
76 pLeu.....SerIleAlaTyrglyArgCysGlnLysIleG 88
|||||::: |||||::: |||||::: |||||::: |||||:::
316 TCTCGTGGGTCCATCACCACACCACTTTGGTATTCTGCCATGATGTG 365
|||||::: |||||::: |||||::: |||||::: |||||:::
88 lyArg 89
|||||
366 GAGCG 370

seq_name: gb_om:ECCOLIIPA

seq_documentation_block:
LOCUS ECCOLIIPA 472 bp mRNA MAM 28-JUL-1995
DEFINITION E.caballus mRNA for colipase A.
ACCESSION X74503
VERSION X74503.1 GI:572676
KEYWORDS anchoring protein; colipase A; secreted protein; triacylglycerol digestion.
SOURCE horse.
ORGANISM Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE 1 (bases 1 to 472)
Kerfelec.B.
AUTHORS Direct Submission
TITLE Submitted (09-AUG-1993) B. Kerfelec, CNRS, CBM, 31 Chemin Joseph
JOURNAL Alquier, 13402 Marseille, Cedex 9, FRANCE
REFERENCE 2 (bases 1 to 472)
Crenon I., Granon S., Chapuis C. and Kerfelec B.

TITLE Molecular cloning and expression of two horse pancreatic cDNA
encoding colipase A and B
JOURNAL Biochim. Biophys. Acta 1213 (3), 357-360 (1994)
MEDLINE 94325330
COMMENT conflicts with previous reported amino acid sequence (Pierrot et al., 1982, Eur.J.Biochem. 123, 347-354:E22->Q22; E32->S32; N89->D89; K92->R92; S/E93->S93; R94->E94; E95.
FEATURES
Location/Qualifiers
source
1..472
/organism="Equus caballus"
/db_xref="taxon:9796"
/clone_lib="cDNA library in lambda gt10"
/tissue_type="pancreas"
<1..33
<1..321
/codon_start=1
/product="colipase A precursor"
/protein_id="CAA52611.1"
/db_xref="GI:572677"
/db_xref="SWISS-PROT:P02704"
/translation="LLILVALAVAYVDPGRGVINLEAGEICLNSACKSECCHOESS
LSLARCAAKASENSECSAWTLGYVYKPCPCERGLTCQVDKTLVGSIMNTNFGICFDA
RSEE"
mat_peptide
34..318
/genes="colipase A"
/product="colipase A"
34..318
/gene="colipase A"
misc_feature
196..210
/genes="colipase A"
/note="interfacial recognition site"
3'UTR
319..471
polyA_signal
441..447
BASE COUNT 99 a 158 c 123 g 92 t
ORIGIN
alignment_scores
Quality: 131.00 Length: 63
Ratio: 2.977 Gaps: 0
Percent Similarity: 69.841 Percent Identity: 38.095
alignment_block
US-09-215-435-179_COPY_24_121 x ECCOLIIPA ..
Align seg 1/1 to: ECCOLIIPA from: 1 to: 472
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||:||||: |||:||||: ||
64 TTGGAGCGGGCGAGATCTGCTGAACAGCGCCAGTGCAAGAGGAATG 113
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 43
|||||:||||: |||:||||: |||:||||: |||:||||: |||
114 CTGCCACAGGAAGAACGAGCTGAGCTGGCGCTGTGCAGCAAGGCCA 163
43 erGluGlySerLeuGluThrGlnValPheGlyGlnTyrArgAla 59
|||||:||||: |||:||||: |||:||||: |||:||||: |||
164 GCGAGACAGCAGAGTGCTGCTGGACACTCTACGGGTTTACTACAG 213
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLys 72
||||| |||:||||: |||:||||: |||:||||: |||:||||: |||
214 TGTCCCTGTGAGCGGGCGCTGACCTGCCAGGTGGACAAG 252
seq_name: gb_ro2:RATCOLIP
seq_documentation_block:
LOCUS RATCOLIP 446 bp mRNA ROD 27-APR-1993
DEFINITION Rat colipase mRNA, complete cds.
ACCESSION M33333
VERSION M33333.1 GI:203502
KEYWORDS colipase.
SOURCE Rattus norvegicus (strain Wistar) (library: puc) pancreas cDNA to mRNA.

ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 446)
AUTHORS Wicker,C. and Puigserver,A.
TITLE Rat pancreatic colipase mRNA: Nucleotide sequence of a cDNA clone
and nutritional regulation by a lipidic diet
JOURNAL Biochem. Biophys. Res. Commun. 167, 130-136 (1990)
MEDLINE 90179738
FEATURES
Location/Qualifiers
source
1..446
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="pancreas"
/tissue_lib="puc"
<1..446
/note="colipase mRNA"
18..356
/codon_start=1
/product="colipase"
/protein_id="AAA40943.1"
/db_xref="GI:203503"
/translation="MKVLVLLVTLVAVAYVPGPGLFINLEGEICVNSMOCKSR
COHDTILGIARCTHKAMENSECSPTLYGIYRCPERGLTCEGDRSIIIGAITNTNYG
VCLDSTRSQ"
sig_peptide
18..68
mat_peptide
69..353
/product="colipase"
polyA_signal
424..430
polyA_site
446
BASE COUNT 96 a 141 c 113 g 96 t
ORIGIN
alignment_scores
Quality: 122.00 Length: 91
Ratio: 2.392 Gaps: 2
Percent Similarity: 56.044 Percent Identity: 31.868
alignment_block
US-09-215-435-179_COPY_24_121 x RATCOLIP ..
Align seg 1/1 to: RATCOLIP from: 1 to: 446
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||:||||: |||:||||: |||:||||: |||
99 CTGGAGGACGGTGAGATCTGCGTAAACAGTATGCAGTGAAGAGCAGATG 148
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 43
|||||:||||: |||:||||: |||:||||: |||:||||: |||
149 CTGCCAACATGACACATCTCTGGGATCGCCGGTGCAACACAGGCCCA 198
43 erGluGlySerLeuGluThrGlnValPheGlyGlnTyrArgAla 59
|||||:||||: |||:||||: |||:||||: |||:||||: |||
199 TGGAGACAGCAGTGCTGCCCAAGACCCCTCTATGGGATCTACTACAG 248
60 CysProCysLeuArgAsnLeuThrCys..... 68
249 TGTCCCTGTGAGCGGGCGCTGACCTGTGAGGGGACAGGAGCATCATGG 298
69IleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||:||||: |||:||||: |||:||||: |||:||||: |||
299 CCCCATCACCAACCAAC.....TAGGGCGTCT 327
84 ysGlnLysIleGlyArgGlnLys 91
||| |||:||||: |||:||||: |||:||||: |||
328 GCTCGACTCCACCGCTCCAAG 350
seq_name: gb_ro2:RATCOLQ
seq_documentation_block:

LOCUS	RATCOLQ	492 bp	mRNA	ROD	04-AUG-1994							
DEFINITION	Rat colipase mRNA, complete cds.											
ACCESSION	M59370											
VERSION	M59370.1	GI:203504										
KEYWORDS	colipase.											
SOURCE	Rat, cDNA to mRNA.											
ORGANISM	Rattus norvegicus											
REFERENCE	Payne, R.M., Sims, H.F., Jennens, M.L. and Lowe, M.E. Rat pancreatic lipase and two related proteins: enzymatic properties and mRNA expression during development. Am. J. Physiol. 266 (5 Pt 1), G914-G921 (1994)											
AUTHORS	Payne, R.M., Sims, H.F., Jennens, M.L. and Lowe, M.E.											
TITLE	Rat pancreatic lipase and two related proteins: enzymatic properties and mRNA expression during development											
JOURNAL	Am. J. Physiol. 266 (5 Pt 1), G914-G921 (1994)											
MEDLINE	94262798											
FEATURES	<p>Location/Qualifiers</p> <p>1..492</p> <p>/organism="Rattus norvegicus"</p> <p>/db_xref="taxon:10116"</p> <p>59..397</p> <p>/product="colipase"</p> <p>/protein_id="AA20505.1"</p> <p>/db_xref="GI:203503"</p> <p>/translation="MKVLVLLVTLVAVAAPGPRGLFINLEDGEICVNSMOCKSRCQHDILIGIAECTHKAMENSECPKTLGIYIYPCPGLTCEGDRSIIGAITNTNYGVCLDSTRSQ"</p>											
BASE COUNT	107 a	163 c	119 g	103 t								
ORIGIN												
alignment_scores:	Quality: 122.00	Length: 91										
	Ratio: 2.392	Gaps: 2										
Percent Similarity:	56.044	Percent Identity:	31.868									
alignment_block:	US-09-215-435-179_COPY_24_121 x RATCOLQ ..											
Align seg 1/1 to: RATCOLQ	from: 1 to: 492											
10	LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy	26										
140	CTGGAGGCGGTGAGATCTGCTGTAACAGATGTCAGTGAAGACAGATG	189										
26	sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS	43										
190	CTGCCAACATGACACCATCTCGGCATCGCCCGGTGCACACAGGCCA	239										
43	erCluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla	59										
240	TGAGAGACACGAGTGTCTCCCAAGACCCCTCTATGGATCTACTACAG	289										
60	CysProCysLeuArgAsnLeuThrCys.....	68										
290	TGTCCTCTGTGAGCGGCGCTGACCTGTGAGGGGACAGGACATCATGG	339										
69IleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC	84										
340	CGCATCATCAACACCAAC.....TACGGCTCT	368										
84	ysGlnLysIleGlyArgGlnLys	91										
369	GCCTCGACTCCACCGCTCCAAG	391										
seq_name:	gb_in3:CEW01B6											
seq_documentation_block:												
LOCUS	CEW01B6	28122 bp	DNA	INV	27-OCT-2000							
DEFINITION	Caenorhabditis elegans cosmid w01B6, complete sequence.											
ACCESSION	268301											

```

NVEGYKFSRDFIDIREKEHDKPRNDHNGSRSLMNRNNEAGRKILKRRKPKCKCHGV
SGACNMKTKRMQPSRFEQVGKILRNKYDKAIRVQINDRNGSLQDLADEATKERKTRALP
TDLVFMDSFDYDQFEDHSGTGLTEGVRKRGSGGAGCDLSCCGRGYNTQEVRSK
CNCKFEWCCKVQVQTCNNVTQVDICK"
complement(join(11401..11572,11702..11769,11814..12098,
12646..13053,13098..13463,13533..13574))
/gene="w01b6.3"
complement(join(11401..11572,11702..11769,11814..12098,
12646..13053,13098..13463,13533..13574))
/gene="w01b6.3"
/notes="Similarity to E.coli tetracycline resistance
protein (SW:TCR4_ECOLI)"
/codon_start=1
/protein_id="CAA92626.1"
/db_xref="GI:3880391"
/db_xref="SPTREMBL:Q23104"
/translacion="MSIYFTSTWQYLSDETATVDFEGAILALQALASALTNPFGCY
WQVSSSTRKPIFCAGISAIQNVLYAFAYLAPPVKKVWMLFARFMTGSPFGALGVLUR
SFITASTQDRMKAVSITNAGFTAGTFLGTQICFIPLGKDGFIQSGSVFLNNMYSIA
AVFMTASIFSILLTFYLEENVYGIISKEDTSDPYFVLPKDFDLFLVLLFYMWMLM
CGVVCVSMAPVFTIAMFNMSREDVLYNGIOTISCFITFIVNFSLASTRLKTIIDNR
LLLLGLLIFEVFIHMPUPFPGPLDRDPTMNSTTLLIGRDFDWCDDTTPRVPLPL
YLFSSIVIGLGPLISSTSSLLSQIIGPRKQTVQGFAPFGSFSQFVVSLESTRL
FELSGKWIIMYHWCIVAVFSVILLWRRLVPLKIPTTNGEATRYKLGTFYRN"
complement(join(14566..14703,14747..14834,14876..15195,
15241..15588,15968..16115,16164..16222))
/gene="w01b6.2"
complement(join(14566..14703,14747..14834,14876..15195,
15241..15588,15968..16115,16164..16222))
/gene="w01b6.2"
/notes="contains similarity to Pfam domain: PF00069
(Eukaryotic protein kinase domain), Score=85.0,
E-value=5.5e-23, N=1
CDNA EST EMBL:AW057393 comes from this gene
CDNA EST EMBL:AW057450 comes from this gene
CDNA EST EMBL:AW057476 comes from this gene"
/codon_start=1
/protein_id="CAA92625.1"
/db_xref="GI:3880390"
/db_xref="SPTREMBL:Q23103"
/translacion="MPSEKEDDYMKNVTFAGKKKFGDWIKGTIDEGGFKVYIATSI
SDPKVAALKAESNEIEGSAIKLEAMILNKANGPVPHIPVHLCAKRLKCYVMY
TLGRNLKLSKLVNNGFSGRTSGRIGICLYALKYVHDNGFIHRDYKPNQFLIG
NETDSERARVTHLDLFGAFPAFVFAHREKNKWTARRARGTAEGRTLYTSPNVHLRK
EOGRDDVDWLLXVYIELNGKALPQWTDQSOREVQOMKLNPAKVVMSPNACMDKV
MPLHASLDYQORDPYHMFKCFQVMENEKITPSSKYDWEKDEPKSVPPASWENPAG
PFQSDPLGINAPPTPEAALTEKSGKKHKK"
16591..16782
/gene="w01b6.4"
16591..16782
/gene="w01b6.4"
/notes="predicted using Genefinder"
/codon_start=1
/protein_id="CAA92622.1"
/db_xref="GI:3880387"
/db_xref="SPTREMBL:Q23101"
/translacion="MCLVQATGGSSTHKLVSAGAEMKIITRPDVPDPGSKDISITRT
ARAPKEDKLVHFTCTRCN"
join(17456..17665,17912..18275,18438..18601,18645..18966,
19023..19175,19227..19510,19569..19682,19730..19780)
/gene="w01b6.5"
join(17456..17665,17912..18275,18438..18601,18645..18966,
19023..19175,19227..19510,19569..19682,19730..19780)
/gene="w01b6.5"
/notes="predicted using Genefinder
Similarity to Mouse tyrosine-protein kinase
(SW:LXN_MOUSE), contains similarity to Pfam domain:
PF00069 (Eukaryotic protein kinase domain), Score=164.4,
E-value=6.1e-46, N=1"
/codon_start=1
/protein_id="CAA92623.1"
/db_xref="GI:3880388"
/db_xref="SPTREMBL:Q23102"

```

```

/translacion="MCTQDEECSPAGDGKQKEKEEHYVVPGRGKEEKKPKGTKPSPIIS
PAFKTKPERNCCQTKNETSGKRLNSNDSPDYKDLQLSEHSGPSTCKIDFQWLEED
EYTKMPCFHGFIGREDLMSLVKNVGFILIRTSVQPNKHEVEKMKQADVKVGLNLSR
EKAPEKKEKEELAGMGDVCRRREVIISVCKDKDEKPADKPOVTPTRNLVIRKDNMG
VHEPLRKFTLTFFCYQKNSGICKETDQFLTPIPLSNWEFIHEDIALQOQKLG
GARGEVRIGMKKLAKSTKKTVEVAVKMLRNAEVVIREQVGLLHEARVHRMDHKNLVR
SYGIADUKLEPLIIMTECAGALREYLRENOETVTLAEKFFVVGSSRGVGYLHKSQKT
IHRDLAVRNLLIHKRHPATINSTVSMGTFLFQPKISDFGLAKISERYENKECKIP
VRLAPLTLESFTFTTKTDFVFCGVITWEIYENGQPHDGKNAQTIIRNLTKKNOFLKL
TNSAPSELRLKLIBERVFTSDPENRCSMTTIVQCAEKTEKPPVGM"
join(22357..22615,22773..22963,23009..23092,23136..23432,
23478..23732,23904..23965,24007..24100)
/gene="w01b6.6"
join(22357..22615,22773..22963,23009..23092,23136..23432,
23478..23732,23904..23965,24007..24100)
/gene="w01b6.6"
/notes="contains similarity to Pfam domain: PF00102
(Protein-tyrosine phosphatase), Score=172.6,
E-value=2e-48, N=1"
/codon_start=1
/protein_id="CAA92627.1"
/db_xref="GI:3880392"
/db_xref="SPTREMBL:Q23105"
/translacion="MSKSTRSRKKKSKGGDRSRKTNKSKRKKKGAKSIDIGKKKKKN
ETVSKDQNTGVSTGAEVKEKEQWSDGDETAKKLVASGAFNSTITSGFKELSNPK
PSDMCSIFKNPNPKSRAPDWPIADNKLILSHAPGNFICAAKVVDVPEFNFTMIVTQI
PDVTSANIEDFWRMIFQEEIVSIVIAVPLECSVTLOQLPILLSGTFSNRGMFLNN
KVDVSAVMTAYTLEVLDPGCSNLSLSTVYHLHNKKOKGLENGELVTVTEKVLKTN
ENTVLSMNGIGRAGTMTLTFTSMLSVQKSKEVNPKEILLKRGERSGLVENAQDPT
VHRAMLWFKNKSDSEEVQKKVFEFANLITTEAVERSSEYVLSQHLIPLILLESMLH
QHQMKVILLKKVRSMSRSL"
complement(join(24622..24664,24714..24833,24928..25089,
25317..25504))
/gene="w01b6.8"
complement(join(24622..24664,24714..24833,24928..25089,
25317..25504))
/gene="w01b6.8"
/notes="predicted using Genefinder"
/codon_start=1
/protein_id="CAA92621.1"
/db_xref="GI:3880386"
/db_xref="SPTREMBL:Q23100"
/translacion="MFNFPFKFPFISQYFNITSEGCSSKIMCLYSQISILFLFWS
LMLQIIGTTCVSKNSQKRIDSLKVPRAPOAQRTNSRISNNQVSLRNSQKSETES
AKTEIINVPLGKAADETSSNLTLPANKQKDLTHTATPKAQAFADTLKRVRRATKPD
DPDEKTSYIP"
complement(join(25616..26241,26289..26568))
/gene="col-2"
/notes="w01b6.7"
complement(join(25616..26241,26289..26568))

```

alignment_scores:

Quality:	90.00	Length:	75
Ratio:	1.957	Gaps:	5
Percent Similarity:	61.333	Percent Identity:	34.667

alignment_block:

```

US-09-215-435-179_copy_24_121 x CEW01B6 ..
Align seg 1/1 to: CEW01B6 from: 1 to: 28122

14 GluSerCysGluArgAsnGlnAspCysGlu.....Th 24
|||||
13131 GAATCTGCGATTGTAATGAACTGCTTTCATCCGATCTTCTGTGTAC 13180
|||||

24 rGlyCysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluL 41
|||||
13181 TTGCTGTACCGATAAAATCTCTGAGAACTCCGAGACGACGAGCTGAAA 13230
|||||

41 ysgGlySerGluGlySerLeuCysGlnThrGlnValPheGlyGlnTyr 57
::: |||||
13231 CCAGTCATGAATCT.....TGCAAACAGCATTTACCAATG 13274
|||||

```


OM of: US-09-215-435-179_COPY_24_121 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:28 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DRV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXP
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179_COPY_24_121

Query length: 98

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAV7614			547.00	1120.56	1.9e-54
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ2674			547.00	1118.23	2.6e-54
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ51465			500.00	1025.77	3.6e-49
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ89828			135.00	265.03	8.6e-07
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ8843			82.50	156.78	8.8e-07
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ7909			82.50	155.57	1.08
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAV75134			82.50	155.57	1.08
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ56830			82.50	154.65	1.21
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ8846			82.50	154.53	1.23
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ8841			82.50	154.49	1.24
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ8842			82.50	154.49	1.24
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ5139			82.50	154.37	1.25
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ56828			82.50	153.24	1.45
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ37042			82.50	140.12	7.81
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ54250			82.50	140.12	7.81
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ5138			82.50	138.54	9.56
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44777			82.00	162.53	0.4410
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ37622			79.50	143.88	4.82
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ74198			78.00	149.08	2.47
/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ87844			77.50	146.53	3.43
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ84529			76.00	138.72	9.35
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35367			76.00	135.03	15.00
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ52675			76.00	126.47	44.95
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ28470			76.00	124.30	59.41
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ28469			76.00	123.15	68.82
/cgnl_9/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ37621			75.50	143.46	5.09
/cgnl_9/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ568221			75.00	131.87	22.48
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ63754			75.00	121.42	85.95
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ55560			75.00	120.52	96.36
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ63721			75.00	120.52	96.36
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ76945			75.00	120.52	96.36
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ25249			75.00	120.52	96.36
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ16991			74.00	142.34	5.87
/cgnl_9/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ58222			74.00	129.69	29.74
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ30093			74.00	122.80	71.93
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ94993			74.00	122.80	71.93
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ94986			74.00	120.12	101.44
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ00679			73.50	140.59	7.35

/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAV7908 + 73.00 127.91 37.38
/cgnl_9/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ52674 + 73.00 117.98 133.63
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ30095 + 73.00 117.98 133.63
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ30094 + 73.00 117.89 135.02
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ7187 - 73.00 116.19 168.04

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ7614

seq_documentation_block:

ID: AAX97614 standard; DNA; 542 BP.

XX AC AAX97614;

DT 13-SEP-1999 (first entry)

DE Extended human secreted protein coding sequence, SEQ ID NO. 78.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.

XX OS Homo sapiens.

XX PN W09931236-A2.

XX PD 24-JUN-1999.

XX PF 17-DEC-1998; 98WO-IB02122.

XX PR 10-AUG-1998; 98US-00961116.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0081563.

XX (GSEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

DR WPI; 1999-385906/32.

DR P-PSDB; AAY35930.

XX New isolated human secreted proteins

PS Claim 1; Page 212; 516pp; English.

XX This sequence represents an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.

SQ Sequence 542 BP; 132 A; 146 C; 126 G; 138 T; 0 other;

alignment_scores:

Quality: 547.00 Length: 98
Ratio: 5.582 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-179_COPY_24_121 x AAX97614

Align seg 1/1 to: AAX97614 from: 1 to: 542

```
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; neurotropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 680 BP; 163 A; 185 C; 167 G; 156 T; 9 other;

alignment_scores:
Quality: 547.00 Length: 98
Ratio: 5.582 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_24_121 x AAA26408
Align seg 1/1 to: AAA26408 from: 1 to: 680

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysI1 17
|||||
202 TCACCTTCTCCCAACAAAATACACCTTTTGGAGCTCAAGGAGTCTTGCA 251

17 eargAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
252 CCGGAACACGAGACTCGGAGACTGGCTGCTGCCACGCTGCTCCAGACAAT 301

34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
|||||
302 GCGAGTCGCACCTCGCGGAGAGAGGGTCCGAGGGCAGTCTGTGTCAAACG 351

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
352 CAGGTGTCTTTGGCCACGATAGAGGCTGCTCCCTGCTGCGGAACCTGAC 401

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
402 TTGTATATATTCAAGAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 451

84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
452 GTCAGAAAATGGAGGCGAGAGCTGGCTAAGAAAATGTTCTTC 495

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AA51465
seq_documentation_block:
ID AAX51465 standard; cDNA; 421 BP.
XX
AC AAX51465;
XX
DT 21-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO:44.
XX
```

```
1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysI1 17
|||||
85 TCACCTTCTCCCAACAAAATATACCTTTTGGAGCTCAAGGAGTCTTGCA 134

17 eargAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
135 CCGGAACACGAGACTCGGAGACTGGCTGCTGCCAAGCTGCTCCAGACAAT 184

34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
|||||
185 GCGAGTCGCACCTCGCGGAGAGGGTCCGAGGGCAGTCTGTGTCAAACG 234

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
235 CAGGTGTCTTTGGCCACATATAGAGCGTGTCCCTGCTGCGGAACCTGAC 284

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
285 TTGTATATATTCAAGAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 334

84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
335 GTCAGAAAATGGAGGCGAGAGCTGGCTAAGAAAATGTTCTTC 378

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA26408
seq_documentation_block:
ID AAA26408 standard; cDNA; 680 BP.
XX
AC AAA26408;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein gene 63 SEQ ID NO:73.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; neurotropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour; ss.
XX
OS Homo sapiens.
XX
FN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PP 29-JUL-1999; 99WO-US17130.
XX
PR 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR P-PSDB; AAY91513.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1; Page 416; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
```


KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; hematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
XX WO9906549-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01231.
XX
XX 01-AUG-1997; 97US-0905279.
XX
XX (GEST) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153779/13.
XX P-PSDB; AAY12687.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
XX Claim 1; Page 177; 522pp; English.
XX
XX AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12681 to
XX AAY12913, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 421 BP; 96 A; 106 C; 110 G; 102 T; 7 other;

alignment_scores:
Quality: 500.00 Length: 95
Ratio: 5.435 Gaps: 0
Percent Similarity: 96.842 Percent Identity: 94.737
alignment_block:
US-09-215-435-179_COPY_24_121 x AAX51465 ...
Align seg 1/1 to: AAX51465 from: 1 to: 421
1 SerLeuSerProThrIlystyrAsnLeuLeuGluLeuLysGluSerCysI1 17
131 TCACITTTCTCCAAACAAATACAACTTTGGAGCTCAAGGACKSKKGCAT 180
17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
181 SGGGAACACGAGGACTCGAGACTGGCTGCTCCCAAGCTGCTCCAGACAATT 230
34 yscluserHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
231 GCGAGTCGCACATCGCCGGAGAGGGGTCGAGGGGAGTCTGTGTCAAACG 280
51 GlnValPhePheGlyGlnIlyrArgAlaCysProCysLeuArgAsnLeuTh 67

|||||
281 CAGGTGTTCTTTGGCCAATATAGAGCGTGTCCCTGCTGCGGAACCTGAC 330
67 rCysIleTyrSerIlyAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
331 TTGTATATATTCAAGAATGAGAAATGGCTTAGCATCGCTATGCCGTT 380
84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLys 95
|||||
381 GTCAGAAATTTGGAAGCGCAGAAAGTTGGCTAGRAAA 415
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ41401
seq_documentation_block:
ID AAZ41401 standard; cDNA; 656 BP.
XX
AC AAZ41401;
XX
XX 19-JAN-2000 (first entry)
XX
XX Human normal pancreas tissue derived cDNA 16.
XX
XX Human; pancreas; cancer; treatment; anticancer; cytostatic;
XX gene therapy; EST; expressed sequence tag; ss.
XX
XX Homo sapiens.
XX
XX DEL9818598-A1.
XX
XX 21-OCT-1999.
XX
XX 19-APR-1998; 98DE-1018598.
XX
XX 19-APR-1998; 98DE-1018598.
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-592039/51.
XX
XX New nucleic acid sequences expressed in normal pancreatic tissues, and
XX derived polypeptides, for treatment of pancreatic cancer and
XX identification of therapeutic agents -
XX
XX Claim 3; Page 56; 92pp; German.

This invention describes novel cDNA sequences (A) that are highly expressed in normal human pancreatic tissue and which have anticancer and cytostatic activity. (A) are used (i) for recombinant expression of polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to identify agents suitable for treatment of pancreatic cancer; (ii) directly for treating this form of cancer (including expression from gene therapy vectors) and (iii) for generation of specific antibodies. (A) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed, so should reduce the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a particular tissue. AAZ41386-241423 represent human normal pancreatic tissue derived cDNA fragments which encode the protein fragments represented in AAY59893-Y59920.

XX Sequence 656 BP; 165 A; 200 C; 162 G; 129 T; 0 other;

alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184
alignment_block:

US-09-215-435-179_COPY_24_121 x AAC241401 ..
Align seg 1/1 to: AAC241401 from: 1 to: 656
10 LeuGluLeuLysGluSerCysLeuArgAsnGlnAspCysGluThrGlyCys 26
170 CTGGAGAACGGGTGAGCTCTGCATGAATAGTGCCTGATGAGAGCAATTG 219
26 SCysGlnArgAlaProAspAsnGluSerHisCysAlaGluLysGlyS 43
220 CTGCCAGCATTCAGTGCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 269
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
270 GCGAGAACGCGAGTCTCTGTCAAGACGCTCTATGGGATTACTACAA 319
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
320 TGTCCCTGTGAGCGTGGCTGACCTGT.....GAGGGAGACAGAC 360
76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 88
361 CATCGTGGCTCCATCACCACCACTTTGGCATCTGCCATGAGCGTG 410
88 lyArgGlnLys 91
411 GACGCTCCAAG 421

seq_name: /cgnl_9/gcgdata/geneseq/NA2000.DAT: AAC98928
seq_documentation_block:
ID: AAC98928 standard; cDNA; 666 BP.
XX AAC98928;
AC
XX
DT 09-MAR-2001 (first entry)
DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:156.
XX Human;
KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neoptotic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative; ss.
XX
OS Homo sapiens.
XX WO200055320-A1.
PN
XX
PD 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05989.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2000-579444/54.
XX P-PSDB; AAB54163.
DR
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition
XX
PS Claim 1; Page 614; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC

neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
gynaecological, cardiant and antiinflammatory activities, and can be used
in gene therapy. The polynucleotide and proteins can be used for
preventing, treating, or ameliorating a medical condition or in assays
for diagnosing a pathological condition or a susceptibility to one in a
subject. Binding partners to the proteins and the activity of the
proteins can be identified. The pancreatic cancer antigens can be used to
detect, treat or prevent pancreatic disorders, especially cancer.
Agonists and antagonists to the antigens can be screened for. The
pancreatic cancer antigen polynucleotides can be used to design nucleic
acid hybridisation probes that can be used in chromosome mapping, linkage
analysis, tissue identification and/or typing and a variety of forensic
and diagnostic methods. The proteins can be used to generate antibodies
which are used to purify, detect and target the polypeptides, including
both in vivo and in vitro diagnostic and therapeutic methods. The
proteins can be used to treat or prevent neural, immune system, muscular,
reproductive, gastrointestinal, pulmonary, cardiovacular, renal or
proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
sequences used in the exemplification of the present invention.
XX
SQ Sequence 666 BP; 151 A; 217 C; 157 G; 133 T; 8 other;

alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184

alignment_block:
US-09-215-435-179_COPY_24_121 x AAC98928 ..

Align seg 1/1 to: AAC98928 from: 1 to: 666

10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCys 26
163 CTGGAGAACGGGTGAGCTCTGCATGAATAGTGCCTGATGAGAGCAATTG 212
26 SCysGlnArgAlaProAspAsnGluSerHisCysAlaGluLysGlyS 43
213 CTGCCAGCATTCAGTGCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 262
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
263 GCGAGAACGCGAGTCTCTGTCAAGACGCTCTATGGGATTACTACAA 312
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
313 TGTCCCTGTGAGCGTGGCTGACCTGT.....GAGGGAGACAGAC 353
76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 88
354 CATCGTGGCTCCATCACCACCACTTTGGCATCTGCCATGAGCGTG 403
88 lyArgGlnLys 91
404 GACGCTCCAAG 414

seq_name: /cgnl_9/gcgdata/geneseq/NA2000.DAT: AAC08843

seq_documentation_block:
ID: AAC08843 standard; DNA; 624 BP.

XX
XX AC AAC08843;
XX
XX DT 01-AUG-2000 (first entry)
XX Human DKR-2 gene splice variant, DKR-2a.
DE
XX DKR-2a; dkk-1; dickkopf-1; antagonist; wnt-8 signaling; morphogenesis;
XX growth factor; cytostatic; sonic hedgehog; tissue differentiation; ss.
KW
XX Homo sapiens.
OS
XX

```
PN WO200018914-A2.
XX
XX
XX 06-APR-2000.
XX
XX 17-SEP-1999; 99WO-US21647.
XX
XX 25-SEP-1998; 98US-0161241.
XX
XX (AMGE-) AMGEN INC.
XX
XX Bass MB, Sullivan JK, Theill LE, Wang D;
XX WPI; 2000-293153/25.
XX P-PSDB; AAY92074.
XX
XX New nucleic acid molecule encoding a biologically active DKR
XX polypeptide, useful in treatment of cancer, e.g. mammary tumors and
XX stem cell tumors
XX
XX PS Claim 1; Page 124; 143pp; English.
XX
XX AAA08838-44 are novel mouse and human genes encoding DKR polypeptides.
XX The human DKR-2a gene is a splice variant of DKR-2 and seems to be
XX missing an exon. Human DKR-2 has about 95 percent homology to murine
XX DKR-2. DKR-1 is a human ortholog of dkk-1 (dickkopf-1), a novel gene
XX identified in Xenopus and mouse, purportedly an antagonist of wnt-8
XX signaling. DKR-2, -3 and -4 are each related to DKR-1 by their cysteine
XX pattern. Dkk-1 is also involved in morphogenesis in the developing
XX embryo, and therefore a growth factor, by inference DKR polypeptides
XX are also growth factors. The DKR polypeptides are useful for treating
XX cancer, e.g. mammary tumors, stem cell tumors, or other cancers in which
XX the wnt and/or sonic hedgehog (shh) signal transduction pathways are
XX activated. They can also be used to enhance tissue differentiation,
XX such as bone formation and hematopoietic cell formation.
XX
XX SQ Sequence 624 BP; 150 A; 161 C; 184 G; 129 T; 0 other;

alignment_scores:
  Quality: 82.50 Length: 90
  Ratio: 2.062 Gaps: 5
  Percent Similarity: 44.444 Percent Identity: 25.556

alignment_block:
US-09-215-435-179_COPY_24_121 x AAA08843 ..
Align seg 1/1 to: AAA08843 from: 1 to: 624
14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
385 GACCCCTGCCCTACGATCATCAGACTGCATTCGAGGGTTTGC..... 426

30 aProAspAsnCysGluSerHis.....CysAlaGluLysG 42
   ||| ||| |||
427 .....TGTGCTCGTCATTTCTGGACCAAAATCTGCAACAGTCG 466

42 LySerGluClySerLeuCysGlnThrGlnValPhePheGlyCln..... 56
   ::|||:|||||:|||||:|||||:|||||:|||||:
467 TCCATCAGGGGAAGTCGTACCAACCAACGCAAGAGGGTTCTCATGGG 516

57 .....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
   ::|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
517 CTGGAATTTTCCACCGTGGCACTGTGCGAAGGCGCTGTCTTGC..... 561

70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
   ::||| ||| ::|||:
562 .....AAAGPATGGAAGATGCACCTACCTACCTCCAAAGCCAGAC 601

84 .....CysGlnLysIle 87
   ||||| ||||| |||||
602 TCCATGTGTGTGAGAAAT 621

seq_name: /cgn1_9/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV07909
```

```
seq_documentation_block:
ID AAV07909 standard; cDNA; 702 BP.
XX
XX AAV07909;
XX
XX 18-JAN-1999 (first entry)
XX
XX Human cysteine-rich secreted protein CRSP-4 cDNA.
XX
XX CRSP-4; cysteine-rich secreted protein 4; tumour; cancer; leukaemia;
XX tissue repair; wound healing; infection; Parkinson's disease;
XX Alzheimer's disease; Huntington's chorea; multiple sclerosis;
XX amyotrophic lateral sclerosis; pontine myelinolysis;
XX human immunodeficiency associated myelopathy; bulbar palsy;
XX spinal muscular atrophy; primary lateral sclerosis; poliomyelitis;
XX Razio-Londe syndrome; Charcot-Marie-Tooth disease; therapy;
XX diagnosis; drug screening; human; signal transduction;
XX cell differentiation; cell proliferation; ss.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..540
XX /*tag= a
XX /note= "claimed fragment"
XX
XX WO9846755-A1.
XX
XX 22-OCT-1998.
XX
XX 16-APR-1998; 98WO-US07894.
XX
XX 20-JAN-1998; 98US-0009802.
XX 16-APR-1997; 97US-0843704.
XX 17-APR-1997; 97US-0842898.
XX 15-JAN-1998; 98US-0071589.
XX
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
XX McCarthy SA;
XX
XX WPI; 1998-568730/48.
XX
XX P-PSDB; AAV73019.
XX
XX New isolated cysteine-rich secreted proteins - used to develop
XX products for treating, e.g. hyperproliferative disorders, cancers,
XX wounds, infectious lesions, degenerative lesions or demyelating
XX diseases
XX
XX Claim 1; Page 104-105; 142pp; English.
XX
XX This partial cDNA clone, the coding region for which is also
XX claimed, codes for novel human cysteine-rich secreted protein 4
XX (CRSP-4, see AAV73019). A partial sequence was derived from dbEST
XX clone AAW55979 using human CRSP-1 as probe. This clone was
XX subsequently obtained from the IMAGE collection and sequenced fully
XX to define the CRSP-4 sequence. CRSP-4 mRNA is expressed in several
XX adult human tissues, with highest expression in hear, brain,
XX placenta, lung and skeletal muscle. CRSP-1, CRSP-2 CRSP-3 and
XX CRSP-4 nucleic acid molecules (see AAV07906-09) and polypeptides (see
XX AAV73016-19) are claimed. The CRSPs have at least 1 of the following
XX activities: (i) modulation of cellular signal transduction, either
XX in vitro or in vivo (e.g. antagonism of the activity of members of
XX the wnt family of secreted proteins or suppression of wnt-dependent
XX signal transduction); (ii) regulation of communication between cells
XX (e.g. regulation of wnt-dependent cell-cell interactions); (iii)
XX regulation of expression of genes whose expression is modulated by
XX binding of CRSP (e.g. CRSP-1) to a receptor; (iv) regulation of gene
XX transcription in a cell involved in development or differentiation;
XX either in vitro or in vivo (e.g. induction of cell differentiation);
XX (v) regulation of cellular proliferation, either in vitro or in vivo
XX (e.g. induction of cellular proliferation or inhibition of
```

CC proliferation as in the case of suppression of tumorigenesis (e.g.
CC suppression of glioblastoma formation)); (vi) formation and
CC maintenance of ordered spatial arrangements of differentiated
CC tissues in vertebrates, both adult and embryonic (e.g. induction of
CC head formation during vertebrate development or maintenance of
CC haematopoietic progenitor cells); (vii) modulation of cell death,
CC such as stimulation of cell survival; (ix) regulation of cell
CC migration, and/or (xii) immune modulation. Products of the
CC invention can be used for e.g. treating hyperproliferative diseases
CC such as neoplastic and hyperplastic disease, e.g. various cancers
CC and leukaemias, and fibroproliferative disorders, malignancy,
CC preneoplasia and benign conditions, to stimulate tissue repair or
CC wound healing or treat spinal cord injuries, brain injuries, lesions
CC associated with surgery, ischemic lesions, malignant lesions,
CC infectious lesions, degenerative lesions (Parkinson's disease,
CC Alzheimer's disease, Huntington's chorea, amyotrophic lateral
CC sclerosis), demyelinating diseases (multiple sclerosis, human
CC immunodeficiency associated myelopathy, pontine myelinolysis),
CC motor neuron injuries, progressive spinal muscular atrophy, and
CC progressive bulbar palsy, primary lateral sclerosis, infantile and
CC juvenile muscular atrophy, progressive bulbar paralysis of childhood
CC (Fazio-Londe syndrome), poliomyelitis, and hereditary motor sensory
CC neuropathy (Charcot-Marie-Tooth disease). The products can also be
CC used for detection, diagnosis and drug screening.
XX
SQ Sequence 702 BP; 221 A; 158 C; 171 G; 152 T; 0 other;

alignment_scores:
Quality: 82.50 Length: 90
Ratio: 2.062 Gaps: 5
Percent Similarity: 44.444 Percent Identity: 25.556

alignment_block:
US-09-215-435-179_COPY_24_121 x AA07909 ..
Align seg 1/1 to: AA07909 from: 1 to: 702

14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30
::: |||||:::||||| ||| |||
301 GACCCCTGCCCTACGATCATCAGACTGCTGGAAGGGGTTTGC..... 342
30 aProAspAsnCysGluSerHis.....CysAlaGluLysG 42
||| ||| |||
343TGTCGCTGCTATTCTGCACCAAAATCTGCAACACCATGTC 382
42 LySerGluGlySerLeuCysGlnThrGlnValPheGlyGln..... 56
:::|||||:::||||| ||| |||
363 TCCATCAGGGGGAGTCTGTACCAAAACGCAAGAGGGTTCTCATGGG 432
57TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
:::|||||:::||||| ||| |||
433 CTGGAAATTTTCCAGCGTTGCGACTGTGCGAAGGGCCCTGCTTGC..... 477
70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
:::|||||:::||||| ||| |||
478AAAGTATGGAAGATGCCACCTACTCTCCAAAGCCAGAC 517
84CysGlnLysIle 87
||||| ||| |||
518 TCCATGTGTGTGCAGAAATTT 537

seq_name: /cgnl_9/gcdata/geneseq/geneseq/NA2000.DAT:AAA75134

seq_documentation_block:

ID AAA75134 standard; cDNA; 702 BP.
XX
AC AAA75134;
XX
DT 15-JAN-2001 (first entry)
XX cDNA encoding a partial human Dickkopf (Dkk)-2 protein.
DE
XX

Human; Dickkopf-2 protein; Dkk-2 protein; Soggy protein; optic disorder;
cysteine-rich secreted protein; glaucoma; conjunctivitis; brain disorder;
Alzheimer's disease; epilepsy; amnesia; inflammation; pulmonary disorder;
skeletal muscle disorder; Goodpasture's syndrome;
cardiovascular disorder; hyperproliferative disorder; cancer; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..540
FT /*tag= a
ET /product= " Dickkopf (Dkk)-2 protein"
XX
WO200052047-A2.
XX
PD 08-SEP-2000.
XX
PF 03-MAR-2000; 2000WO-US05452.
XX
PR 05-MAR-1999; 99US-0263022.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy S;
XX
DR WPI: 2000-579276/54.
DR P-PSDB; AAB08877.
XX
XX Human Dickkopf (hDkk) and Soggy nucleic acids and proteins, useful as
modulating agents in regulating cellular processes and particularly for
treating disorders characterized by aberrant expression or activity of
Dkk, e.g. Alzheimer's,
XX
PS Example 6; Page 81-82; 208pp; English.
XX
CC The present sequence encodes a partial Dickkopf (Dkk)-2 protein. The
CC specification also describes Soggy (Dkk-related) sequences. Dkk is a
CC cysteine-rich secreted protein. The Dkk nucleic acids and proteins are
CC useful as modulating agents in regulating cellular processes. They are
CC particularly useful in treating subjects having a disorder characterized
CC by aberrant expression or activity of Dkk such as optic disorders
CC (glaucoma, conjunctivitis), brain disorders (Alzheimer's disease,
CC epilepsy, amnesia), inflammation, skeletal muscle disorders, pulmonary
CC disorders (Goodpasture's syndrome), cardiovascular disorders, and
CC hyperproliferative disorders (cancer). The Dkk proteins and nucleic
CC acids may also be used for research purposes, such as for chromosome
CC mapping, tissue typing and in screening assays to identify modulators.
XX
SQ Sequence 702 BP; 221 A; 158 C; 171 G; 152 T; 0 other;

alignment_scores:
Quality: 82.50 Length: 90
Ratio: 2.062 Gaps: 5
Percent Similarity: 44.444 Percent Identity: 25.556
alignment_block:
US-09-215-435-179_COPY_24_121 x AAA75134 ..
Align seg 1/1 to: AAA75134 from: 1 to: 702
14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30
::: |||||:::||||| ||| |||
301 GACCCCTGCCCTACGATCATCAGACTGCTGGAAGGGGTTTGC..... 342
30 aProAspAsnCysGluSerHis.....CysAlaGluLysG 42
||| ||| |||
343TGTCGCTGCTATTCTGCACCAAAATCTGCAACACCATGTC 382
42 LySerGluGlySerLeuCysGlnThrGlnValPheGlyGln..... 56
:::|||||:::||||| ||| |||
383 TCCATCAGGGGGAGTCTGTACCAAAACGCAAGAGGGTTCTCATGGG 432

57TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
 433 CTGGAATTTTCAGCGTTGCGACTGTGCGAAGGGCGCTGCTTGC..... 477
 70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
 478AAAGTATGGAAGATGCCACTACTCTCTCCAAAGCCAGAC 517
 84CysGlnLysIle 87
 518 TCCATGTGTGTGAGAAAT 537

seq_name: /cgml_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AAx56830

seq_documentation_block:

ID AAX56830 standard; cDNA: 768 BP.

XX AAX56830;

DT 14-JUL-1999 (first entry)

DE Human phdtk-2 cDNA.

XX Signal pathway; wnt; inhibitor; secreted glycoprotein; receptor;
 KW therapy; diagnosis; treatment; cancer; breast; carcinoma; melanoma;
 KW colon; cell proliferation; differentiation; ss.

OS Homo sapiens.

XX WO9922000-A1.

PD 06-MAY-1999.

PF 27-OCT-1998; 98WO-DE03155.

XX 27-OCT-1997; 97DE-1047418.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PA Glinka A, Niehrs C;

XX WPI; 1999-303017/25.

XX Protein that inhibits wnt signalling

PS Claim 4a; Fig 2; 39pp; German.

XX This invention describes an inhibitory protein of the wnt signalling
 CC pathway (which comprises secreted glycoproteins and their associated
 CC receptors). The products of the invention are useful as therapeutic
 CC and diagnostic agents, e.g. to treat cancer (e.g. carcinoma of breast
 CC and colon, or melanoma) and for studying processes associated with wnt
 CC signalling (cell proliferation and differentiation). Primers derived
 CC from the nucleic acid of the invention are used to detect the expression
 CC of the gene. AAX56827-X26833 are sequences used in the method of the
 CC invention.

SQ Sequence 768 BP; 239 A; 172 C; 186 G; 171 T; 0 other;

alignment_scores:

Quality:	82.50	Length:	90
Ratio:	2.062	Gaps:	5
Percent Similarity:	44.444	Percent Identity:	25.556

alignment_block:

US-09-215-435-179_COPY_24_121 x AAX56830 ..

Align seq 1/1 to: AAX56830 from: 1 to: 768

14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30

339 GACCCCTGCGTCAGCATCATCAGACTGCATTGAAGGGTTTTCG..... 380

30 aProAspAsnCysGluSerHis.....CysAlaGluLysG 42
 381TGTGCTCGTCATTTCTGGACCAAAATCTGCARACCACTGC 420
 42 lySerGluGlySerLeuCysGlnThrGlnValPheGlyGln..... 56
 421 TCATCATCAGGGGAAGTCTGTACCAACAACGCAAGAAGGGTTCTCATGGG 470
 57TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
 471 CTGGAAATTTTCAGCGTTGCGACTGTGCGAAGGGCGCTGCTTGC..... 515
 70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
 516AAAGTATGGAAGATGCCACTACTCTCTCCAAAGCCAGAC 555
 84CysGlnLysIle 87
 556 TCCATGTGTGTGAGAAAT 575

seq_name: /cgml_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA08846

seq_documentation_block:

ID AAA08846 standard; DNA: 777 BP.

XX AAA08846;

DT 01-AUG-2000 (first entry)

DE Human DKR-2 DNA optimized for expression in E. coli.

KW DKR-2; recombinant; bacterial expression; dkk-1; dickkopf-1; antagonist;
 KW wnt-8 signalling; morphogenesis; growth factor; cytotstatic;
 KW sonic hedgehog; tissue differentiation; ss.

XX Synthetic.

OS Homo sapiens.

XX WO200018914-A2.

XX 06-APR-2000.

PF 17-SEP-1999; 99WO-US21647.

XX 25-SEP-1998; 98US-0161241.

XX (AMGE-) AMGEN INC.

XX Bass MB, Sullivan JK, Theill LE, Wang D;

XX WPI; 2000-293153/25.

XX New nucleic acid molecule encoding a biologically active DKR
 PT polypeptide, useful in treatment of cancer, e.g. mammary tumors and
 PT stem cell tumors

XX Claim 1; Page 142-143; 143pp; English.

XX AAX08845-48 are recombinant human DKR constructs with codons optimized
 CC for expression in Escherichia coli. Recombinant DKR-2 was engineered to
 CC encode a 235 amino acid polypeptide.

XX DKR-1 is a human ortholog of dkk-1 (dickkopf-1), a novel gene

CC identified in Xenopus and mouse, purportedly an antagonist of wnt-8
 CC signaling. DKR-2, -3 and -4 are each related to DKR-1 by their cysteine
 CC pattern. Dkk-1 is also involved in morphogenesis in the developing
 CC embryo, and therefore a growth factor, by inference DKR polypeptides
 CC are also growth factors. The DKR polypeptides are useful for treating
 CC cancer, e.g. mammary tumors, stem cell tumors, or other cancers in which
 CC the wnt and/or sonic hedgehog (shh) signal transduction pathways are
 CC activated. They can also be used to enhance tissue differentiation,
 CC such as bone formation and hematopoietic cell formation.

XX	Mus sp.		
OS			
XX	key	Location/Qualifiers	
FH	sig_peptide	1..99	
FT		/*tag= a	
FT	mat_peptide	100..778	
FT		/*tag= b	
FT			
XX	WQ200018914-A2.		
XX			
XX	06-APR-2000.		
PD			
XX			
PF	17-SEP-1999;	99WO-US21647.	
XX			
PR	25-SEP-1998;	98US-0161241.	
XX			
XX	(AMGE-) AMGEN INC.		
PA			
XX			
PI	Bass MB, Sullivan JK, Theill LE, Wang D;		
XX			
XX	WPI; 2000-293153/25.		
DR	P-PSDB; AAY92072.		
DR			

XX	Key	Location/Qualifiers
FH		

```
FT sig_peptide 1..99
FT /*tag= a
FT mat_peptide 100..778
FT /*tag= b
XX
XX WO200018914-A2.
XX
XX 06-APR-2000.
XX
XX 17-SEP-1999; 99WO-US21647.
XX
XX 25-SEP-1998; 98US-0161241.
XX
XX (AMGE-) AMGEN INC.
XX
XX Bass MB, Sullivan JK, Theill LE, Wang D;
XX
XX WPI: 2000-293153/25.
XX
XX P-PSDB; AAY92073.
XX
XX New nucleic acid molecule encoding a biologically active DKR
XX polypeptide, useful in treatment of cancer, e.g. mammary tumors and
XX stem cell tumors
XX
XX Claim 1; Page 122-123; 143pp; English.
XX
XX AAA08838-44 are novel mouse and human genes encoding DKR polypeptides.
XX The human DKR-2 gene has about 95 percent homology to murine DKR-2.
XX DKR-1 is a human ortholog of dkk-1 (dickkopf-1), a novel gene identified
XX in Xenopus and mouse, purportedly an antagonist of wnt-8 signaling.
XX DKR-2, -3 and -4 are each related to DKR-1 by their cysteine pattern.
XX DKR-1 is also involved in morphogenesis in the developing embryo, and
XX therefore a growth factor, by inference DKR polypeptides are also
XX growth factors. The DKR polypeptides are useful for treating cancer,
XX e.g. mammary tumors, stem cell tumors, or other cancers in which the wnt
XX and/or sonic hedgehog (shh) signal transduction pathways are activated.
XX They can also be used to enhance tissue differentiation, such as bone
XX formation and hematopoietic cell formation.
XX
XX Sequence 780 BP; 200 A; 203 C; 215 G; 162 T; 0 other;

alignment_scores:
    Quality: 82.50      Length: 90
    Ratio: 2.062       Gaps: 5
    Percent Similarity: 44.444      Percent Identity: 25.556

alignment_block:
US-09-215-435-179_COPY_24_121 x AAA08842 ..

Align seg 1/1 to: AAA08842 from: 1 to: 780

14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30
   :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GACCCCTGCCTACGATCATGACACTGCATTTGAAGGGTTTGC..... 582
   30 aproAspAsnCysGluSerHis.....CysAlaGluLysG 42
583 .....TGTGTCGTCATTTCTGGACCAAAATCTGCACCAACAGTCG 622
   42 lySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 56
   ::|||:|||||:|||||:|||||:|||||:|||||:
623 TCATCATCGGGGGAAGTCTGTACCAACAACGCAAGAAGGGTTCTCATGG 672
   57 .....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
   :::::|||||:|||||:|||||:|||||:|||||:
673 CTGGAATTTTCACGCGTTGCGACTGTGCGAAGGGCCTGCTCTGC..... 717
   70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
   :::::|||||:|||||:|||||:|||||:
718 .....AAAGTATGGAAGATGCGACCTACTCTCTCCAAAGCCAGAC 757
   84 .....CysGlnLysIle 87
```

```
|||||
758 TCCATGTGTGTCAGAAAT 777

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA75139

seq_documentation_block:
ID AAA75139 standard; cDNA; 789 BP.
XX
XX AAA75139;
XX
XX 15-JAN-2001 (first entry)
XX
XX Open reading frame encoding a human Dickkopf (Dkk)-2 protein.
XX
XX Human; Dickkopf-2 protein; Dkk-2 protein; Soggy protein; optic disorder;
XX cysteine-rich secreted protein; glaucoma; conjunctivitis; brain disorder;
XX Alzheimer's disease; epilepsy; amnesia; inflammation; pulmonary disorder;
XX skeletal muscle disorder; Goodpasture's syndrome;
XX cardiovascular disorder; hyperproliferative disorder; cancer; ss.
XX
XX Homo sapiens.
XX
XX OS
XX FH
XX Key Location/Qualifiers
XX CDS 1..789
XX FT /*tag= a
XX FT /product= " Dickkopf (Dkk)-2 protein"
XX FT sig_peptide 1..102
XX FT /*tag= b
XX
XX WO200052047-A2.
XX
XX 08-SEP-2000.
XX
XX 03-MAR-2000; 2000WO-US05452.
XX
XX 05-MAR-1999; 99US-0263022.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy S;
XX
XX WPI: 2000-579276/54.
XX
XX P-PSDB; AAB08880.
XX
XX Human Dickkopf (hDkk) and Soggy nucleic acids and proteins, useful as
XX modulating agents in regulating cellular processes and particularly for
XX treating disorders characterized by aberrant expression or activity of
XX Dkk, e.g. Alzheimer's, -
XX
XX Claim 2; Page 97-98; 208pp; English.
XX
XX The present sequence encodes a human Dickkopf (Dkk)-2 protein. The
XX specification also describes Soggy (Dkk-related) sequences. Dkk is a
XX cysteine-rich secreted protein. The Dkk nucleic acids and proteins are
XX useful as modulating agents in regulating cellular processes. They are
XX particularly useful in treating subjects having a disorder characterized
XX by aberrant expression or activity of Dkk such as optic disorders
XX (glaucoma, conjunctivitis), brain disorders (Alzheimer's disease,
XX epilepsy, amnesia), inflammation, skeletal muscle disorders, pulmonary
XX disorders (Goodpasture's syndrome), cardiovascular disorders, and
XX hyperproliferative disorders (cancer). The Dkk proteins and nucleic
XX acids may also be used for research purposes, such as for chromosome
XX mapping, tissue typing and in screening assays to identify modulators.
XX
XX Sequence 789 BP; 200 A; 208 C; 219 G; 162 T; 0 other;

alignment_scores:
    Quality: 82.50      Length: 90
    Ratio: 2.062       Gaps: 5
    Percent Similarity: 44.444      Percent Identity: 25.556

alignment_block:
```

XX	DE	Mouse pmdkk-2 cDNA.	seq_name: /cgnl_9/cgcdata/geneseq/Na2000.DAT:AAA37042
XX	DE	Signal pathway; wnt; inhibitor; secreted glycoprotein; receptor;	seq_documentation_block:
KW	KW	therapy; diagnosis; treatment; cancer; breast; carcinoma; melanoma;	ID AAA37042 standard; cDNA; 3170 BP.
KW	KW	colon; cell proliferation; differentiation; ss.	AC AAA37042;
XX	OS	Mus sp.	XX 08-AUG-2000 (first entry)
XX	XX	WO9922000-A1.	XX Human PRO1316 (UNQ682) cDNA sequence SEQ ID NO:69.
PN	XX	06-MAY-1999.	XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
PD	XX	27-OCT-1998; 98WO-DE03155.	KW transmembrane; secretion; Immunoadhesion; pharmaceutical; screening;
XX	PF	27-OCT-1997; 97DE-1047418.	KW ss.
XX	XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	OS Homo sapiens.
PA	XX	Glinka A, Niehrs C;	XX WO200012708-A2.
PI	XX	WPI; 1999-303017/25.	XX 09-MAR-2000.
XX	XX	Protein that inhibits wnt signalling	XX 01-SEP-1999; 99WO-US20111.
XX	PT	Claim 4a; Fig 2; 39pp; German.	XX 01-SEP-1998; 98US-0098716.
XX	PS	This invention describes an inhibitory protein of the wnt signalling	XX 01-SEP-1998; 98US-0098749.
XX	CC	pathway (which comprises secreted glycoproteins and their associated	XX 01-SEP-1998; 98US-0098750.
XX	CC	receptors). The products of the invention are useful as therapeutic	XX 02-SEP-1998; 98US-0098803.
XX	CC	and diagnostic agents, e.g. to treat cancer (e.g. carcinoma of breast	XX 02-SEP-1998; 98US-0098821.
XX	CC	and colon, or melanoma) and for studying processes associated with wnt	XX 02-SEP-1998; 98US-0098843.
XX	CC	signalling (cell proliferation and differentiation). Primers derived	XX 09-SEP-1998; 98US-0099536.
XX	CC	from the nucleic acid of the invention are used to detect the expression	XX 09-SEP-1998; 98US-0099596.
XX	CC	of the gene. AAX56827-X26833 are sequences used in the method of the	XX 09-SEP-1998; 98US-0099602.
XX	CC	invention.	XX 09-SEP-1998; 98US-0099642.
XX	CC		XX 10-SEP-1998; 98US-0099741.
XX	CC		XX 10-SEP-1998; 98US-0099754.


```
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 15-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103349.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.

PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106484.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
PR XX
PR XX
PA (GETH ) GENENTECH INC.
PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
DR P-PSDB; AAY99360.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PT small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2; Fig 41; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
SQ Sequence 3170 BP; 1014 A; 613 C; 629 G; 914 T; 0 other;

alignment_scores:
    Quality: 82.50      Length: 90
    Ratio: 2.062      Gaps: 5
    Percent Similarity: 44.444      Percent Identity: 25.556

alignment_block:
    US-09-215-435-179_COPY_24_121 x AAA37042      ..
Align seg 1/1 to: AAA37042 from: 1 to: 3170
```

Thu Nov 15 10:52:05 2001

```
14 GluSerCysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAl 30
   :: |||::|||::|||::|||::||| ||| |||
751 GACCCCTGCCTACGATCATCAGACTGCATTGAAGGTTTTCG,..... 792
30 aProAspAsnCysGluSerHis.....CysAlaGluLysG 42
   ||| |||
793 .....TGTGCTCGTCATTCTGTGGACCAAAATCTGCAAAACCAGTGC 832
42 lySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 56
   ::|||::|||::|||::||| ||| |||::|
833 TCCATCAGGGGGAAGTCTGTACCAACAACGCAAGAAGGTTCTCATGGG 882
57 .....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTy 70
   ::||| ||| ||| ::|||::|||::|||
883 CTGGAATAATTTCACAGCGTTCCGACTGTGGAGGCGCTGTCTTGC.... 927
70 rSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArg..... 83
   ::||| ||| ::|||::|||::|
928 .....AAAGTATGGAAAGATGCCACCTACTCTCCAAAGCCAGAC 967
84 .....CysGlnLysIle 87
   |||::|||::|||
968 TCCATGTGTCTCAGAAAATT 987
```

OM of: US-09-215-435-179_COPY_24_121 to: Issued_Patents_NA:* out_format : pfs

Date: Nov 15, 2001 4:11 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_1411001_1411004_2141/app_query.fasta_1.1519
-DB-Issued_Patents_NA -QPMF-fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMF=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@cgnl_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179_COPY_24_121

Query length: 98

Database: Issued_Patents_NA.*

Database sequences: 351203

Database length: 113238999

Search time (sec): 228.030000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2 +		86.50	79.17	9.2e+03	4403765
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-11 +		76.00	133.39	8.84	3226
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-313-288B-11 +		76.00	133.39	8.84	3226
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-03164-11 +		76.00	133.39	8.84	3226
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-597-545-3 +		75.00	139.01	4.30	1556
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-457-135-3 +		75.00	139.01	4.30	1556
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-882-046-3 +		75.00	128.05	17.55	4315
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-181-706-1 +		75.00	127.11	19.78	4707
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-458-791-1 +		75.00	127.11	19.78	4707
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-459-066-1 +		75.00	127.11	19.78	4707
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-597-545-4 +		74.00	136.67	5.81	1573
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-457-135-4 +		74.00	136.67	5.81	1573
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-9 +		73.00	124.34	28.23	4029
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-313-288B-9 +		73.00	124.34	28.23	4029
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-03164-9 +		73.00	124.34	28.23	4029
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-085-199B-3 -		73.00	122.47	35.89	4796
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-551-437-4 +		70.50	127.84	18.02	1734
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-004-225-4 +		70.50	127.84	18.02	1734
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-084-346-4 +		70.50	127.84	18.02	1734
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-104-704-4 +		70.50	127.84	18.02	1734
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-214-278-4 +		70.50	118.99	56.11	3955
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-400-159-7 +		70.50	117.69	66.29	4464
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-611-729A-7 +		70.50	117.64	66.68	4483
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-924-063A-4 -		70.00	142.60	2.72	396
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-085-199B-1 -		69.00	128.79	15.95	1164
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-188-930-255 +		69.00	126.33	21.88	1464
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-188-930-73 +		69.00	125.16	25.43	1633
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-13 +		68.00	121.79	39.15	1816
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-313-288B-13 +		68.00	121.79	39.15	1816
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-03164-13 +		68.00	121.79	39.15	1816
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-083-590A-21 +		67.50	102.66	455.55	9723
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-532-384-21 +		67.50	102.66	455.55	9723
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-226-264-27 +		67.00	130.06	13.56	684
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-454-455-3 +		67.00	114.28	102.61	2972
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-454-455-5 +		67.00	111.67	143.39	3789
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-537-361E-7 -		66.50	115.56	87.02	2378
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-525-940-22 +		66.00	114.37	101.46	2397
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-976-838-22 +		66.00	114.37	101.46	2397
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-525-940-20 +		66.00	113.32	116.07	2643
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-976-838-20 +		66.00	113.32	116.07	2643

/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-525-940-19 + 66.00 112.91 122.29
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-976-838-19 + 66.00 112.91 122.29
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-525-940-17 + 66.00 112.83 123.58
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-976-838-17 + 66.00 112.83 123.58
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-976-838-28 - 66.00 112.83 123.58

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2

seq_documentation_block:

; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

alignment_scores:

Quality: 86.50 Length: 73
Ratio: 2.403 Gaps: 5
Percent Similarity: 49.315 Percent Identity: 30.137

alignment_block:

US-09-215-435-179_COPY_24_121 x US-09-103-840A-2 ..
Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765
16 CysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProas 32
3963587TGTGTCGGAACCGAGTGTGCCAATGCCGTGTGTC...CGAAGCCCC... 3963631
32 pAsnCysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysG 49
3963632.AGTTGCTGCTCCGCGTGTCCGGCGC... 3963658
49 InThrGlnValPhePheGlyGlnTyrArgAlaCysProCys..... 62
3963659.....CGATGTTGCGCTGCCGCGCTTC 3963682
63 ...LeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLeu 78
3963683CGGTGCGGACGAGTGTGTGTCGCGGTATCCGAACCCGACGTTGG...AG 3963729

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-11

seq_documentation_block:

; Sequence 11, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

us-09-215-435-179_copy_24_121.rni

Thu Nov 15 10:52:06 2001

```

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136...2543
; US-07-862-021B-11

alignment_scores:
  Quality: 76.00 Length: 101
  Ratio: 1.407 Gaps: 4
  Percent Similarity: 53.465 Percent Identity: 21.782

alignment_block:
  US-09-215-435-179_COPY_24_121 x US-07-862-021B-11 ..
  Align seg 1/1 to: US-07-862-021B-11 from: 1 to: 3226

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
1747 GAGAACTGAGAATGTTATGTAATGAGGAATGCTCCCTAGCAGCTG 1796

27 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAlaG 40
1797 CTTTGTCCAGCAATGGGAGAGTGGGATGAATGCAATGCTAGCTGTG 1846

40 lu.....LysGlySerGluGly 45
1847 CAGGAATGAAAGCGCACAGATGATCAAGATGACTCTCTGCTGATGGA 1896

46 SerLeuLysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 62
1897 TCTATGTCGAGGCGAGAACTACAGAGCGAGAAATGCATGATGCCCGA 1946

62 sLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 77
1947 ATGCCACTATATCCCTCTTCTATCCCTCCCTGCTGCTGCTGCTGCTG 1996

77 euSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 93
1997 GCAGCGTGACATGTGGGAGGGAATGCCAACC...CGCAAAAGGATGCTG 2043

```

```

94 Lys 94
1004 AAA 2046
seq_name: /cgnl_7/ptodata/1/lna/5A_COMB.seq:US-08-313-288B-11
seq_documentation_block:
; Sequence 11, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136..2543
; US-08-313-288B-11

```

```

alignment_scores:
  Quality: 76.00 Length: 101
  Ratio: 1.407 Gaps: 4
  Percent Similarity: 53.465 Percent Identity: 21.782

alignment_block:
  US-09-215-435-179_COPY_24_121 x US-08-313-288B-11 ..
  Align seg 1/1 to: US-08-313-288B-11 from: 1 to: 3226

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
1747 GAGAACTGAGAATGTTATGTAATGAGGAATGCTCCCTAGCAGCTG 1796

27 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAlaG 40
1797 CTTTGTCCAGCAATGGGAGAGTGGGATGAATGCAATGCTAGCTGTG 1846

40 lu.....LysGlySerGluGly 45
1847 CAGGAATGAAAGCGCACAGATGATCAAGATGACTCTCTGCTGATGGA 1896

```


ENTER REFLECTION DATA:
; APPLICATION NUMBER: --to be assigned-- (USSN 08/958,598

```

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/458,791
; FILING DATE: 10-Dec-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE: 28-OCT-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; LENGTH: 4707 base pairs
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
;
; US-09-458-791-1
;
;
; alignment_scores:
;   Quality: 75.00      Length: 78
;   Ratio: 1.705      Gaps: 7
;   Percent Similarity: 56.410      Percent Identity: 33.333
;
; alignment_block:
; US-09-215-435-179_COPY_24_121 x US-09-458-791-1 ..
;
; Align seg 1/1 to: US-09-458-791-1 from: 1 to: 4707
;
; 33 AsnCysGlu.....SerHisCysAlaGluLysGlySerG1 44
;   ::::: |||  ::::: |||  ::::: |||  ::::: |||
; 1296 TCCTGTGAAGAATATCTACATTTATCTAACAGCTGGGAAGAGGTGAGGA 1345
;
; 44 uGlySerLeuCysGlnThrGln.....ValPhePheGlyGlnTyr. 57
;   ::::: |||  ::::: |||  ::::: |||  ::::: |||
; 1346 GAATTCGTGTGCAAACTGCAATAAACATAAAATCCCTTCGGAGTGTTTA 1395
;
; 58 .....ArgAla..CysProCysLeuArgAsnLeuTh 67
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1396 ACAGCCACAGACCCCTCAGTGGTGGTCCCATTCGCTACAAAGGTGCAC 1445
;
; 67 r.....CysIleTyrSerLysAsn...GluLysTTrpLeuSerI 79
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1446 TTTTCAAGGAGATTGTGTACATTTCAGAGAACTTAGAAAACTGGCTGGATA 1495
;
; 79 leAlaTyrGly.....ArgCysGlnLysIle 87
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1496 TTTTCGTCTGGAGCAAAAAGTGCCTAAAATT 1527
;
; seq_name: /cgnl_7/ptodata/1/ina/68_COMB.seq:US-09-459-066-1
;
; seq_documentation_block:
; ; Sequence 1, Application US/09459066
; ; Patent No. 6187909
; ; GENERAL INFORMATION:
; ; APPLICANT: SPRIGGS, Melanie
; ; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; ; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
;
;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/459,066
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
;
; US-09-459-066-1
;
;
; alignment_scores:
;   Quality: 75.00      Length: 78
;   Ratio: 1.705      Gaps: 7
;   Percent Similarity: 56.410      Percent Identity: 33.333
;
; alignment_block:
; US-09-215-435-179_COPY_24_121 x US-09-459-066-1 ..
;
; Align seg 1/1 to: US-09-459-066-1 from: 1 to: 4707
;
; 33 AsnCysGlu.....SerHisCysAlaGluLysGlySerG1 44
;   ::::: |||  ::::: |||  ::::: |||  ::::: |||
; 1296 TCCTGTGAAGAATATCTACATTTATCTAACAGCTGGGAAGAGGTGAGGA 1345
;
; 44 uGlySerLeuCysGlnThrGln.....ValPhePheGlyGlnTyr. 57
;   ::::: |||  ::::: |||  ::::: |||  ::::: |||
; 1346 GAATTCGTGTGCAAACTGCAATAAACATAAAATCCCTTCGGAGTGTTTA 1395
;
; 58 .....ArgAla..CysProCysLeuArgAsnLeuTh 67
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1396 ACAGCCACAGACCCCTCAGTGGTGGTCCCATTCGCTACAAAGGTGCAC 1445
;
; 67 r.....CysIleTyrSerLysAsn...GluLysTTrpLeuSerI 79
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1446 TTTTCAAGGAGATTGTGTACATTTCAGAGAACTTAGAAAACTGGCTGGATA 1495
;
; 79 leAlaTyrGly.....ArgCysGlnLysIle 87
;   |||  |||  |||  |||  |||  |||  |||  |||
; 1496 TTTTCGTCTGGAGCAAAAAGTGCCTAAAATT 1527
;
; seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-597-545-4
```



```
seq_documentation_block:
; Sequence 4, Application US/08597545
; Patent No. 5580738
; GENERAL INFORMATION:
; APPLICANT: LABORDA, Jorge
; TITLE OF INVENTION: Delta-Like Gene Expressed In
; TITLE OF INVENTION: Neuroendocrine Tumors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,545
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/989,537
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1573 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; STRAIN: Mouse Dlk
; US-08-597-545-4

alignment_scores:
Quality: 74.00 Length: 61
Ratio: 2.176 Gaps: 4
Percent Similarity: 55.738 Percent Identity: 31.148

alignment_block:
US-09-215-435-179_copy_24_121 x US-08-597-545-4 ..
Align seg 1/1 to: US-08-597-545-4 from: 1 to: 1573
12 LeuLysGluSerCysTleArgAsnGlnAspCysGluThrGlyCysCysG1 28
||| : : : : : : : : : : : : : : : : : : : : : : : : :
290 C|C|T|G|T|G|C|A|A|G|T|G|T|A|C|T|G|C|C|C|T|G|G|T|G|T|C|A|T|G|G|A|G|C|T|G|C|.. 337
28 n|A|G|A|P|A|P|A|P|A|S|n|C|y|S|G|l|U|S|e|r|H|I|S|C|y|S|A|A|G|l|U|L|y|S|G|l|y|S|e|r|G|l|U|G| 45
: : : : : : : : : : : : : : : : : : : : : : : : :
338 .A|A|G|A|A|C|C|A|T|G|C|A|G|T|G|C|.....AT|C|T|C|A|A|G|A|T|G|G|C|T|G|G|G|A|C|G| 377
45 l|y|S|e|r|L|e|u|C|y|S|G|l|n|T|H|r|G|l|n|V|a|l|P|H|e|P|H|e|G|l|n|T|y|r|A|r|G|A|l|a|C|y|S|... 60
| : : : : : : : : : : : : : : : : : : : : : : : : :
378 G|G|A|A|T|T|C|T|G|C|G|A|A|T|A|G|A|C|G|T|.....CG|G|G|C|T|T|G|C|A|C|C| 412
61 .....P|r|o|C|y|S|L|e|u|A|r|G|A|S|n|L|e|u|T|H|r|C|y|S|I|l|e| 69
413 T|C|A|A|C|C|C|C|T|G|C|C|C|A|C|A|A|T|G|G|A|C|T|T|G|C|G|T|G| 445

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-597-545-4
```

```
seq_documentation_block:
; Sequence 4, Application US/08457135
; Patent No. 5644031
; GENERAL INFORMATION:
; APPLICANT: LABORDA, Jorge
; TITLE OF INVENTION: Delta-Like Gene Expressed In
; TITLE OF INVENTION: Neuroendocrine Tumors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,135
; FILING DATE: 01-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/989,537
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/304/NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1573 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; STRAIN: Mouse Dlk
; US-08-457-135-4

alignment_scores:
Quality: 74.00 Length: 61
Ratio: 2.176 Gaps: 4
Percent Similarity: 55.738 Percent Identity: 31.148

alignment_block:
US-09-215-435-179_copy_24_121 x US-08-457-135-4 ..
Align seg 1/1 to: US-08-457-135-4 from: 1 to: 1573
12 LeuLysGluSerCysTleArgAsnGlnAspCysGluThrGlyCysCysG1 28
||| : : : : : : : : : : : : : : : : : : : : : : : : :
290 C|C|T|G|T|G|C|A|A|G|T|G|T|A|C|T|G|C|C|C|T|G|G|T|G|T|C|A|T|G|G|A|G|C|T|G|C|.. 337
28 n|A|G|A|P|A|P|A|P|A|S|n|C|y|S|G|l|U|S|e|r|H|I|S|C|y|S|A|A|G|l|U|L|y|S|G|l|y|S|e|r|G|l|U|G| 45
: : : : : : : : : : : : : : : : : : : : : : : : :
338 .A|A|G|A|A|C|C|A|T|G|C|A|G|T|G|C|.....AT|C|T|C|A|A|G|A|T|G|G|C|T|G|G|G|A|C|G| 377
45 l|y|S|e|r|L|e|u|C|y|S|G|l|n|T|H|r|G|l|n|V|a|l|P|H|e|P|H|e|G|l|n|T|y|r|A|r|G|A|l|a|C|y|S|... 60
| : : : : : : : : : : : : : : : : : : : : : : : : :
378 G|G|A|A|T|T|C|T|G|C|G|A|A|T|A|G|A|C|G|T|.....CG|G|G|C|T|T|G|C|A|C|C| 412
61 .....P|r|o|C|y|S|L|e|u|A|r|G|A|S|n|L|e|u|T|H|r|C|y|S|I|l|e| 69
413 T|C|A|A|C|C|C|C|T|G|C|C|C|A|C|A|A|T|G|G|A|C|T|T|G|C|G|T|G| 445

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-862-021B-9
```

```

seq_documentation_block:
; Sequence 9, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/862.021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226..2647
; US-07-862-021B-9

alignment_scores:
Quality: 73.00 Length: 103
Ratio: 1.304 Gaps: 4
Percent Similarity: 54.369 Percent Identity: 20.388

alignment_block:
US-09-215-435-179_COPY_24_121 x US-07-862-021B-9 ..
Align seg 1/1 to: US-07-862-021B-9 from: 1 to: 4029

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1852 GAAGAGACAGAGAAGTGCACGGTCAACGAGGAGTGTCTCTCTAGCAGCTG 1901
27 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAla. 39
| : : : : : ||||| ||||| ||||| ||||| |||||
1902 CCTGGTGACTAGTGGGGTGTAGTGGGATGACTGCAGCGCCACCTGTGGAA 1951
40 .....GlutysGlySerGluGly 45
: : : : : ||||| ||||| ||||| ||||| |||||
1952 TGGGCATGAGAGAGCGCACCGCATGGTCAAGATGATGACCCCGGGAGCGC 2001
46 SerLeuCysGlnThrGlnValPhePheGlyClnTyArgAlaCysProCy 62
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2002 TCCATGTGCAGGCGGAGACTTCCAGCGGGAGAAATGCATGATGCTCTGA 2051
62 sLeuArgAsnLeuThrCysIleTySerLysAsnGluLysTrp.....L 77

```

```

2052 GTGCGATACCATCCGCTGTGTCTCTCTCTGTCGAGTGGAGCGACT 2101
      ::::::::::: |||::: ||| ::::::::::|||
77 euSerIleAlaTyrgGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 93
      |||::: |||::: ::::: |||::: |||:::
2102 GTACGGTGACCTGTGGGAGGCGATCGGAGC...CGCCAGCGATGCTC 2148
      |||::: |||::: ::::: |||::: |||:::
94 LysLysMet 96
      |||::: |||:::
2149 AAGTCTCTG 2157
      |||::: |||:::

seq_name: /cgnl_7/ptodata/1/lna/5A_COMB.seq:US-08-313-288B-9

seq_documentation_block:
; Sequence 9, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226..2647
; US-08-313-288B-9

alignment_scores:
      Quality: 73.00 Length: 103
      Ratio: 1.304 Gaps: 4
Percent Similarity: 54.369 Percent Identity: 20.388

alignment_block:
US-09-215-435-179_COPY_24_121 x US-08-313-288B-9 ..

Align seg 1/1 to: US-08-313-288B-9 from: 1 to: 4029

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysC 27
      ||| |||::: |||::: |||::: |||::: |||:::
1852 GAAGAGACAGAGAGTGCACGGTCAACGAGGAGTGCTCTCTAGCAGCTG 1901
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
27 sGlnAraAlaPro.....AspAsnCysGluSerHisCysAla. 39

```

```
1 1902 CCTGTGACTGAGTGGGTGAGTGGGATGACGACGCCACCTGTGGAA 1951
40 40 .....GluLysGlySerGluGly 45
1952 TGGCGATGAGAGCGGACCGCATGGTCAAGATGAGCGCCCGGACGGC 2001
46 SerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 62
2002 TCCATGTGACGAGCGGAGACTTCGACGGGAGAGAAATGCATGATGCCTGA 2051
62 sleuArgasnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 77
2052 GTGCCATACCATCCGCTGCTCTCCTTGGTCCGAGTGGAGCGACT 2101
77 euserIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 93
2102 GTAGCGTGACCTGTGGGAGGCGCATGCCGACG...CGCCAGCGGATGCTC 2148
94 LysLysMet 96
2149 AAGTCTCTG 2157
```

seq_name: /cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US93-03164-9

```
seq_documentation_block:
; Sequence 9, Application PC/TUS9303164
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03164
; FILING DATE: 19930402
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4029 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 226..2647
PCT-US93-03164-9
```

```
alignment_scores:
Quality: 73.00 Length: 103
Ratio: 1.304 Gaps: 4
Percent Similarity: 54.369 Percent Identity: 20.388
alignment_block:
US-09-215-435-179_copy_24_121 x PCT-US93-03164-9 ..
Align seg 1/1 to: PCT-US93-03164-9 from: 1 to: 4029
11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
1852 GAAGACAGAGAGTGCACGTCACAGGAGTGTCTCTCTAGCAGCTG 1901
27 scInArgAlaPro.....AspAsnCysGluSerHisCysAla. 39
1902 CCTGTGACTGAGTGGGTGAGTGGGATGACTGCAGCGCCACTGTGGAA 1951
40 .....GluLysGlySerGluGly 45
1952 TGGCGATGAGAGCGGACCGCATGGTCAAGATGAGCGCCCGGACGGC 2001
46 SerLeuCysGlnThrGlnValPheGlyGlnTyrArgAlaCysProCy 62
2002 TCCATGTGACGAGCGGAGACTTCGACGGGAGAGAAATGCATGATGCCTGA 2051
62 sleuArgasnLeuThrCysIleTyrSerLysAsnGluLysTrp.....L 77
2052 GTGCCATACCATCCGCTGCTCTCCTTGGTCCGAGTGGAGCGACT 2101
77 euserIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAla 93
2102 GTAGCGTGACCTGTGGGAGGCGCATGCCGACG...CGCCAGCGGATGCTC 2148
94 LysLysMet 96
2149 AAGTCTCTG 2157
```

us-09-215-435-179_copy_24_121.rni

Thu Nov 15 10:52:06 2001


```
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
PCT-US99-17130-73

alignment_scores:
    Quality: 547.00      Length: 98
    Ratio: 5.582         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_copy_24_121 x PCT-US99-17130-73  ..
Align seg 1/1 to: PCT-US99-17130-73 from: 1 to: 680

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCys11 17
|||||
202 TCACCTTCTCCAAACAAATACAACTTTTGGAGCTCAAGGAGCTTGCAT 251

17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
252 CCGGAACCCAGGACTCGGAGACTGGCTGCTGCCAACGTGCTCCAGACAATT 301

34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuLysGlnThr 50
|||||
302 GCGAGTCGCAGCTCGCGGAGAGAGGGTCCGAGGCGAGTCTGTGTCAACG 351

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
352 CAGGTGTCTTGGCCAGTATAGAGCGTGTCCCTGCCGGAACCTGCAT 401

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
402 TTGTATATATCAAGAATGAGAAATGGCTTAGCATCGCCTATGCGCGTT 451

84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
452 GTCAGAAAATTGGAGGAGAGAGTTGGCTAAGAAAATGTTCTTC 495

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-489-847-73

seq_documentation_block:
; Sequence 73, Application US/09489847
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454

; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-73

alignment_scores:
    Quality: 547.00      Length: 98
    Ratio: 5.582         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_copy_24_121 x US-09-489-847-73  ..
Align seg 1/1 to: US-09-489-847-73 from: 1 to: 680

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCys11 17
|||||
202 TCACCTTCTCCAAACAAATACAACTTTTGGAGCTCAAGGAGCTTGCAT 251

17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
252 CCGGAACCCAGGACTCGGAGACTGGCTGCTGCCAACGTGCTCCAGACAATT 301

34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuLysGlnThr 50
|||||
302 GCGAGTCGCAGCTCGCGGAGAGAGGGTCCGAGGCGAGTCTGTGTCAACG 351

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
352 CAGGTGTCTTGGCCAGTATAGAGCGTGTCCCTGCCGGAACCTGCAT 401

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
402 TTGTATATATCAAGAATGAGAAATGGCTTAGCATCGCCTATGCGCGTT 451

84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
452 GTCAGAAAATTGGAGGAGAGAGTTGGCTAAGAAAATGTTCTTC 495

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-4817

seq_documentation_block:
; Sequence 4817, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
```

us-09-215-435-179_copy_24_121.rnnp

Thu Nov 15 10:52:07 2001

```

; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Ives
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..461
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.80000019073486
; OTHER INFORMATION: seq FLLFFFLFLTRG/SL
; US-09-834-366-4817

alignment_scores:
    Quality: 542.00      Length: 98
    Ratio: 5.588        Gaps: 0
    Percent Similarity: 98.980      Percent Identity: 98.980

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-834-366-4817 ..
Align seg 1/1 to: US-09-834-366-4817 from: 1 to: 533

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysIl 17
|||||
168 TCACCTTTCTCCAAACAAATACACCTTTTGGAGCTCAAGGAGCTTGCAT 217
|||||
17 eArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnC 34
|||||
218 CCGAACACGAGACTGCGAGACTGCTGCCAACGTCCTCCAGACAATT 267
|||||
34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
|||||
268 GCGAGTCGCACTGCGGAGAGGGTCCGAGGGCAGTCGTCGTCGTCGTC 317
|||||
51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
318 CAGGTGTTCTTTGGCCARTATAGAGCGTGTCCCTGCCTGCGGAACCTGAC 367
|||||
67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
368 TTGTATATATTCAAAGAAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 417
|||||
84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
418 GTCAGAAAATTTGGAAGGCAGAGTTGGCTAARAAAATGTTCTTC 461
|||||

seq_name: /cgnl_7/ptodata/1/pna/us6019_COMB.seq:US-60-197-873-4817

seq_documentation_block:
; Sequence 4817, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Ives
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873

```

```

; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4817
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..461
; NAME/KEY: sig_peptide
; LOCATION: 99..167
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.80000019073486
; OTHER INFORMATION: seq FLLFFFLFLTRG/SL
; US-60-197-873-4817

alignment_scores:
    Quality: 542.00      Length: 98
    Ratio: 5.588        Gaps: 0
    Percent Similarity: 98.980      Percent Identity: 98.980

alignment_block:
US-09-215-435-179_COPY_24_121 x US-60-197-873-4817 ..
Align seg 1/1 to: US-60-197-873-4817 from: 1 to: 533

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysIl 17
|||||
168 TCACCTTTCTCCAAACAAATACACCTTTTGGAGCTCAAGGAGCTTGCAT 217
|||||
17 eArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnC 34
|||||
218 CCGAACACGAGACTGCGAGACTGCTGCCAACGTCCTCCAGACAATT 267
|||||
34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
|||||
268 GCGAGTCGCACTGCGGAGAGGGTCCGAGGGCAGTCGTCGTCGTCGTC 317
|||||
51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
318 CAGGTGTTCTTTGGCCARTATAGAGCGTGTCCCTGCCTGCGGAACCTGAC 367
|||||
67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
368 TTGTATATATTCAAAGAAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 417
|||||
84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
418 GTCAGAAAATTTGGAAGGCAGAGTTGGCTAARAAAATGTTCTTC 461
|||||

seq_name: /cgnl_7/ptodata/1/pna/us081_COMB.seq:US-08-196-481-520

seq_documentation_block:
; Sequence 520, Application US/08196481
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William
; APPLICANT: Rosen, Craig
; APPLICANT: Ruben, Steve
; APPLICANT: Dillon, Patrick
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences, and Expression
; TITLE OF INVENTION: Products
; NUMBER OF SEQUENCES: 1719
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
; ADDRESS: Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068

```



```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot
; REGISTRATION NUMBER: 24, 025
; REFERENCE/DOCKET NUMBER: 325800-92
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-196-481-520

alignment_scores:
  Quality: 532.00      Length: 98
  Ratio: 5.485        Gaps: 0
Percent Similarity: 98.980 Percent Identity: 96.939

alignment_block:
US-09-215-435-179_COPY_24_121 x US-08-196-481-520 ..

Align seg 1/1 to: US-08-196-481-520 from: 1 to: 488

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysII 17
|||||
56 TCACCTTCTCCAAACAAATACAAACCTTTTGGAGCTCAAGGAGCTTTGCAT 105

17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
106 CCGGAACCAAGGAGCTGGAGCTGGCTGCCAACGTCTCCAGACAATT 155

34 ysgluSerHisCysAlaGluLysGlySerGluCysGlnThr 50
|||||
156 GCGAGTCGCACTCGCGGAGAGGGGTCCGAGGCGAGTCTGTGTCAAACG 205

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
206 CAGGTGTTCTTGGNCAATATAGAGCGTGTCCCTGCGGGAACCTGAC 255

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
256 TTGTATATATCAAGAAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 305

84 ysglnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
306 GTCAGAAAATTNGAAGCGAGAAGTTGGTTAAGGAAATGTTCTTT 349

seq_name: /cgnl_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481A-520

seq_documentation_block:
; Sequence 520, Application US/08196481A
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steve
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences and Expression
```

Products

```
; NUMBER OF SEQUENCES: 1721
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,481A
; FILING DATE: February 15, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PO-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-196-481A-520

alignment_scores:
  Quality: 532.00      Length: 98
  Ratio: 5.485        Gaps: 0
Percent Similarity: 98.980 Percent Identity: 96.939

alignment_block:
US-09-215-435-179_COPY_24_121 x US-08-196-481A-520 ..

Align seg 1/1 to: US-08-196-481A-520 from: 1 to: 488

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysII 17
|||||
56 TCACCTTCTCCAAACAAATACAAACCTTTTGGAGCTCAAGGAGCTTTGCAT 105

17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
106 CCGGAACCAAGGAGCTGGAGCTGGCTGCCAACGTCTCCAGACAATT 155

34 ysgluSerHisCysAlaGluLysGlySerGluCysGlnThr 50
|||||
156 GCGAGTCGCACTCGCGGAGAGGGGTCCGAGGCGAGTCTGTGTCAAACG 205

51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
206 CAGGTGTTCTTGGNCAATATAGAGCGTGTCCCTGCGGGAACCTGAC 255

67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
256 TTGTATATATCAAGAAATGAGAAATGGCTTAGCATCGCCTATGGCCGTT 305

84 ysglnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
306 GTCAGAAAATTNGAAGCGAGAAGTTGGTTAAGGAAATGTTCTTT 349

seq_name: /cgnl_7/ptodata/1/pna/US081_COMB.seq:US-08-196-481B-520

seq_documentation_block:
```

```

; Sequence 520, Application US/08196481B
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 7
; FILE REFERENCE: PO07
; CURRENT APPLICATION NUMBER: US/08/196,481B
; CURRENT FILING DATE: 1994-02-15
; NUMBER OF SEQ ID NOS: 1721
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 520
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (220)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (317)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (354)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (356)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (366)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (378)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (425)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (434)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (437)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (440)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (449)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (457)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (458)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (463)
; OTHER INFORMATION: n equals a,t,g, or c

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (482)
; OTHER INFORMATION: n equals a,t,g, or c
US-08-196-481B-520

alignment_scores:
    Quality: 532.00      Length: 98
    Ratio: 5.485        Gaps: 0
    Percent Similarity: 98.980    Percent Identity: 96.939

alignment_block:
US-09-215-435-179_COPY_24_121 x US-08-196-481B-520 ..
Align seg 1/1 to: US-08-196-481B-520 from: 1 to: 488

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCysI1 17
|||||
56 TCACCTTTCTCCACAAATACACCTTTTGGAGCTCAAGGAGTCTTGCAT 105
|||||
17 eArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAspAsnC 34
|||||
106 CCGAACCAGGACTGCCGAGACTGGCTGCTGCCACACGTGCTCCAGACAATT 155
|||||
34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuCysGlnThr 50
|||||
156 GCGAGTCGCACCTCGCGGAGAAAGGGTCCGAGGCGAGTCTGTGTCAAACG 205
|||||
51 GlnValPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
206 CAGGTGTTCTTTGGNCAATATAGAGCGGTGCCCTGCCCTATGCCCTATGCGC 305
|||||
67 cCysIleTyrSerLysAsnGluLysTyrLeuSerIleAlaTyrGlyArgC 84
|||||
256 TTGTATATATTCAGAAGATGAGAAATGCTTAGCATCGCCTATGCGCCTT 305
|||||
84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
|||||
306 GTCAGAAAAATTNGAAGGAGAGAGTGGTTAAGGAAATGTTCTTT 349
|||||

seq_name: /cgnl_7/ptodata/1/pna/US089_COMB.seq:US-08-905-279-44

seq_documentation_block:
; Sequence 44, Application US/08905279
; GENERAL INFORMATION:
; APPLICANT: Genset SA
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; TITLE OF INVENTION: EXPRESSED IN TESTIS AND OTHER TISSUES
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,279
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.011a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176

```

INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 421 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Testis
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 62..130
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 9.8
OTHER INFORMATION: seq FLFFFLFLTRG/SL
US-08-905-279-44

alignment_scores:
Quality: 500.00 Length: 95
Ratio: 5.435 Gaps: 0
Percent Similarity: 96.842 Percent Identity: 94.737

alignment_block:
US-09-215-435-179_COPY_24_121 x US-08-905-279-44 ..

Align seg 1/1 to: US-08-905-279-44 from: 1 to: 421

1 SerLeuSerProThrLysTyrAsnLeuLeuGluLeuLysGluSerCys11 17
|||||
131 TCACTTTCACCAACAAATACAACTTTGGAGCTCAAGGAGSKKGCAT 180
17 eargAsnGlnAspCysGluThrGlyCysGlnArgAlaProAspAsnC 34
|||||
181 SGGGAACCAAGAGCTCGAGACTGGCTGCTGCCAACGTCTCCAGACAAT 230
34 ysGluSerHisCysAlaGluLysGlySerGluGlySerLeuLeuCysGlnThr 50
|||||
231 GCGAGTCGCACCTCGCGGAGAGGGTCCGAGGCGAGTCTGTGTCAAAGC 280
51 GlnValPhePheGlyClnTyrArgAlaCysProCysLeuArgAsnLeuTh 67
|||||
281 CAGGTGTTCTTTGGCAATATAGAGCGTCTCCCTCGCGGAACCTGAC 330
67 rCysIleTyrSerLysAsnGluLysTrpLeuSerIleAlaTyrGlyArgC 84
|||||
331 TTGTATATTCAAGAAATGAGAAATGGCTTAGCATCGCCTATGGCCGPT 380
84 ysGlnLysIleGlyArgGlnLysLeuAlaLysLys 95
|||||
381 GTCAGAAATTCGAGGCGAGAGTGGCTAGRAAA 415

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-184-698-48

seq_documentation_block:

; Sequence 48, Application US/60184698
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M. Jones, Anissa L.
; APPLICANT: Lincoln, Stephen E. Yu, Jimmy Y.
; APPLICANT: Russo, Frank D. Greenawalt, Lila B.
; APPLICANT: Spiro, Peter A. Panzer, Scott R.
; APPLICANT: Banville, Steve C. Roseberry, Ann M.
; APPLICANT: Bratcher, Shawn R. Wright, Rachel J.
; APPLICANT: Dufour, Gerard E. Chen, Wensheng
; APPLICANT: Cohen, Howard J. Liu, Tommy
; APPLICANT: Rosen, Bruce Yap, Pierre E.
; APPLICANT: Shah, Purvi Amshey, Stefan
; APPLICANT: Chalup, Michael S. Fong, Willy Tuen
; TITLE OF INVENTION: BIOCHEMICAL PATHWAY MOLECULES
; FILE REFERENCE: PT-0115 P
; CURRENT APPLICATION NUMBER: US/60/184.698
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 1132

CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 1132
SOFTWARE: PERL Program
SEQ ID NO 48
LENGTH: 463
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 474641.1
FEATURE:
NAME/KEY: unsure
LOCATION: 347-368
OTHER INFORMATION: a, t, c, g, or other
US-60-184-698-48

alignment_scores:
Quality: 497.00 Length: 88
Ratio: 5.648 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-179_COPY_24_121 x US-60-184-698-48 ..

Align seg 1/1 to: US-60-184-698-48 from: 1 to: 463

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysC 27
|||||
61 GAGCTCAAGGAGTCTTGCATCCGGAACCCAGAGCTCGAGACTGGCTG 110
27 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 44
|||||
111 CCACAGTGTCTCCAGACAATGGAGTCGCACCTCGCGGAGAGGGTCCG 160
44 luGlySerLeuCysGlnThrGlnValPhePheGlyClnTyrArgAlaCys 60
|||||
161 AGGGCAGTCTGTCTCAACGCCAGGTCTTCTTGGCCAATATAGAGCTGT 210
61 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 77
|||||
211 CCTCGCTCGGAACCTGACTTGTATATATCAAGAAATGAGAAATGGCT 260
77 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 94
|||||
261 TAGCATCGCTATGGCGTGTGCAGAAATTCGAGGCGAGAGTGGCTA 310
94 ysLysMetPhePhe 98
|||||
311 AGAAATGTTCTTC 324

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-184-698-694

seq_documentation_block:

; Sequence 694, Application US/60184698
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M. Jones, Anissa L.
; APPLICANT: Lincoln, Stephen E. Yu, Jimmy Y.
; APPLICANT: Russo, Frank D. Greenawalt, Lila B.
; APPLICANT: Spiro, Peter A. Panzer, Scott R.
; APPLICANT: Banville, Steve C. Roseberry, Ann M.
; APPLICANT: Bratcher, Shawn R. Wright, Rachel J.
; APPLICANT: Dufour, Gerard E. Chen, Wensheng
; APPLICANT: Cohen, Howard J. Liu, Tommy
; APPLICANT: Rosen, Bruce Yap, Pierre E.
; APPLICANT: Shah, Purvi Amshey, Stefan
; APPLICANT: Chalup, Michael S. Fong, Willy Tuen
; TITLE OF INVENTION: BIOCHEMICAL PATHWAY MOLECULES
; FILE REFERENCE: PT-0115 P
; CURRENT APPLICATION NUMBER: US/60/184.698
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 1132

Thu Nov 15 10:52:07 2001

```
; SOFTWARE: PERL Program
; SEQ ID NO 694
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 474641.1.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 347-368
; OTHER INFORMATION: a, t, c, g, or other
US-60-184-698-694

alignment_scores:
  Quality: 497.00      Length: 88
  Ratio: 5.648        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-179_COPY_24_121 x PCT-US00-31337-1 ..
  Align seg 1/1 to: PCT-US00-31337-1 from: 1 to: 523

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
|||||
91 GAGCTCAAGCAGCTTGGCATCCGGAACCAAGGAGCTGGCTGCTG 140

27 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 44
|||||
141 CCAACGTGCTCCAGACAATTCGAGTGCACATCGCGGAGAGGGGTCCG 190

44 LuGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 60
|||||
191 AGGCAGTCTGTGTCAAACGACGAGTGTCTTTGGCCCAATATAGAGCGTGT 240

61 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 77
|||||
241 CCCTGCGCTCGGAACTGACTTGTATATATTCAAAGAAATGAGAAATGGCT 290

77 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 94
|||||
291 TAGCATCGCCTATGGCGTGTGTGCAAAAATTGGAAAGCAGCAAGTTGGCTA 340

94 ysLysMetPhePhe 98
|||||
341 AGAAAATGTTCTTC 354

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:US-09-443-634-1

seq_documentation_block:
; Sequence 1, Application US/09443634
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Colipase-Like Protein - 1
; FILE REFERENCE: 99-85
; CURRENT APPLICATION NUMBER: US/09/443,634
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(354)
US-09-443-634-1

alignment_scores:
  Quality: 497.00      Length: 88
  Ratio: 5.648        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-215-435-179_COPY_24_121 x US-09-443-634-1 ..
  Align seg 1/1 to: US-09-443-634-1 from: 1 to: 523

11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
|||||
91 GAGCTCAAGCAGCTTGGCATCCGGAACCAAGGAGCTGGCTGCTG 140

27 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 44
|||||
141 CCAACGTGCTCCAGACAATTCGAGTGCACATCGCGGAGAGGGGTCCG 190

44 LuGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 60
|||||
191 AGGCAGTCTGTGTCAAACGACGAGTGTCTTTGGCCCAATATAGAGCGTGT 240

61 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 77
|||||
241 CCCTGCGCTCGGAACTGACTTGTATATATTCAAAGAAATGAGAAATGGCT 290

77 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 94
|||||
291 TAGCATCGCCTATGGCGTGTGTGCAAAAATTGGAAAGCAGCAAGTTGGCTA 340

94 ysLysMetPhePhe 98
|||||
341 AGAAAATGTTCTTC 354

seq_name: /cgnl_7/ptodata/1/pna/PCTUS_COMB.seq:PCT-US00-31337-1

seq_documentation_block:
; Sequence 1, Application PCT/US0031337
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: EDUCATIONAL KIT AND METHOD USING
; TITLE OF INVENTION: COLIPASE-LIKE PROTEIN-1
; FILE REFERENCE: 99-85PC
; CURRENT APPLICATION NUMBER: PCT/US00/31337
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/166,626
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(354)
PCT-US00-31337-1
```

```
141 CCAACGTGCTCCAGACAAATTGGCGAGTCCGACCTCCGCGAGAGGGGTCCG 190
44 luGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 60
|||||
191 AGGCAGCTCTGTGTCAAACGCCAGGTGCTTTGGCCCAATATAGAGCGTGT 240
61 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 77
|||||
241 CCCTGCCTCGGAACCTGACTTGTATATATCAAGAAGATGAGAAATGGCT 290
77 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 94
|||||
291 TAGCATCGCTATGGCGTGTGTGAGAAAATTGGAGGCAGAGTTGGCTA 340
94 ysLysMetPhePhe 98
|||||
341 AGAAAATGTTCTTC 354
```

```
seq_name: /cgnl_7/ptodata/1/pna/US097A_COMB.seq:US-09-713-067-1
seq_documentation_block:
; Sequence 1, Application US/09713067
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: EDUCATIONAL KIT AND METHOD USING
; TITLE OF INVENTION: COLIPASE-LIKE PROTEIN-1
; FILE REFERENCE: 99-85US
; CURRENT APPLICATION NUMBER: US/09/713,067
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/166,626
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(354)
US-09-713-067-1
```

```
alignment_scores:
  Quality: 497.00      Length: 88
  Ratio: 5.648        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:

US-09-215-435-179_COPY_24_121 x US-09-713-067-1 ..

Align seg 1/1 to: US-09-713-067-1 from: 1 to: 523

```
11 GluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
|||||
91 GAGCTCAAGGAGTCTTGCATCCGGAACCAAGGACTGCGAGACTGGCTGCTG 140
27 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 44
|||||
141 CCAACGTGCTCCAGACAAATTGGCGAGTCCGACCTCCGCGAGAGGGGTCCG 190
44 luGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCys 60
|||||
191 AGGCAGCTCTGTGTCAAACGCCAGGTGCTTTGGCCCAATATAGAGCGTGT 240
61 ProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTrpLe 77
|||||
241 CCCTGCCTCGGAACCTGACTTGTATATATCAAGAAGATGAGAAATGGCT 290
77 uSerIleAlaTyrGlyArgCysGlnLysIleGlyArgGlnLysLeuAlaL 94
|||||
291 TAGCATCGCTATGGCGTGTGTGAGAAAATTGGAGGCAGAGTTGGCTA 340
```

```
94 ysLysMetPhePhe 98
|||||
341 AGAAAATGTTCTTC 354
```

us-09-215-435-179_copy_24_121.rnrm

Thu Nov 15 10:52:07 2001

OM of: US-09-215-435-179_COPY_24_121 to: Pending_Patents_NA_New:* out_format : pfs
Date: Nov 15, 2001 4:09 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2194/app_query.fasta_1.1519
-DB=Pending_Patents_NA_New -QPMF=fastap -SUFFIX=rnpn
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-XGAPEXT=0.000 -XGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000
-YGAPEXT=0.500 -YGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09215435_@cgnl_1_69 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-179_COPY_24_121
Query length: 98
Database: Pending_Patents_NA_New:*
Database sequences: 127876
Database length: 122380375
Search time (sec): 564.420000

score_list:	Sequence	Strd Orig	zscore	EScore	Len	Documentation
308	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-703-559-9	+	74.50	136.21	2.26	
452	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-703-559-7	+	74.50	131.88	3.94	
2	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-970-366-206	+	74.00	138.06	1.78	
3	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-970-966-184	+	74.00	135.07	2.62	
3	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-970-966-177	+	74.00	132.10	3.83	
4	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-594-353	+	74.00	129.88	5.09	
3	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-5020	+	71.50	151.57	0.3152	
3	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-5019	+	71.50	150.70	0.3524	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-335-448-2381	-	71.50	114.47	36.72	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-702	+	70.00	133.67	3.13	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-189-455	+	70.00	133.67	3.13	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-192-455	+	70.00	133.67	3.13	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-697-455	+	70.00	133.67	3.13	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-824-455	+	70.00	133.67	3.13	
17	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-974-298-88	+	70.00	131.76	4.00	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-594-1085	+	70.00	131.76	4.00	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1145	+	70.00	126.74	7.61	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-1145	+	69.00	131.15	4.32	
637	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-619-252-1	+	69.00	115.07	34.02	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-8642	+	66.50	139.89	1.41	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-8641	+	66.00	138.77	1.63	
14	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-971-392-43	-	66.00	125.02	9.49	
5	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-594-862	+	66.00	109.21	72.12	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-976-516-351	+	65.50	100.71	214.41	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2006	+	65.50	100.71	214.41	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2007	+	65.50	100.71	214.41	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-973-278-680	+	65.00	125.38	9.07	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-560	+	65.00	118.02	23.30	
2	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-1210	+	64.50	134.85	2.69	
10	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-886-055-28	-	64.50	125.18	9.30	
2	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-518	+	64.50	113.84	39.81	
111	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-960-858-5	+	64.00	96.99	345.92	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-332	-	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-472	+	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-473	+	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2390	-	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2391	-	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2392	+	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2393	+	64.00	93.01	576.29	
1	/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2394	+	64.00	93.01	576.29	

/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2395 + 64.00 93.01 576.29
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2396 - 64.00 93.01 576.29
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2397 + 64.00 93.01 576.29
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-742 + 63.50 122.66 12.85
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-4416 + 63.50 110.44 61.63
seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-703-559-9

seq_documentation_block:
; Sequence 9, Application US/09703559
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Insulin-Responsive Sequence DNA Binding Protein-1 and Methods
; FILE REFERENCE: E056 1010
; CURRENT APPLICATION NUMBER: US/09/703,559
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 3087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3084)
US-09-703-559-9

alignment_scores:
Quality: 74.50 Length: 63
Ratio: 2.328 Gaps: 5
Percent Similarity: 50.794 Percent Identity: 33.333

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-703-559-9 ..
Align seg 1/1 to: US-09-703-559-9 from: 1 to: 3087

21 AspCysGluThrGly.....CysCysGlnAArgAlaProAs 32
205 GAGTGGCCCCGAGGCTTCATGGCCCTGGACTGCGAGGAGAGATCCCCGA 254
32 pAsnCysGlu.....SerHisCysAlaGluLysGlySerGluG 45
255 TGACTGTGAGTCCCGACGAGGAGATGCCTG.....GGCGCCACACA 238
45 lySerLeuCysGlnThrGlnVal...PhePheGly..... 55
299 CCACCCCTCTGCCAGTCCCGCTGGATCTTGGCTTCTCTGTGAATTT 348
56 GlnTyrArgAlaCysProCysLeuArgAsnLeuThrCys 68
349 GAAATCAGCCATGCGCTGCAACATGAACACACACAGTGC 387
seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-703-559-7

seq_documentation_block:
; Sequence 7, Application US/09703559
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Insulin-Responsive Sequence DNA Binding Protein-1 and Methods
; FILE REFERENCE: E056 1010
; CURRENT APPLICATION NUMBER: US/09/703,559
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4524
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

us-09-215-435-179_copy_24_121.rnpn

Thu Nov 15 10:52:08 2001

```
; NAME/KEY: 5'UTR
; LOCATION: (1)..(58)
; NAME/KEY: exon
; LOCATION: (59)..(3082)
; NAME/KEY: 3'UTR
; LOCATION: (3083)..(4524)
US-09-703-559-7

alignment_scores:
  Quality: 74.50      Length: 63
  Ratio: 2.328        Gaps: 5
  Percent Similarity: 50.794  Percent Identity: 33.333

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-703-559-7 ..
Align seg 1/1 to: US-09-703-559-7 from: 1 to: 4524

21 AspCysGluThrGly.....CysCysGlnArgAlaProAs 32
   ::::|
263 GAGTGGCCCCGAGGCTTCATGGCCCTGGACTCGAGGAGAGAGTCCCCGA 312
   ::::|
32 pAsnCysGlu.....SerHisCysAlaGluLysGlySerGluG 45
   ::::|
313 TGACTGTGAGTGGCGGCAACGAGGAGGAGATGCTCTG.....GGCGCAACA 356
   ::::|
45 lySerLeuCysGlnThrGlnVal....PhePheGly..... 55
   ::::|
357 CCACCCCTCGCCAGTGGCCCTCGGATCTTTGGGCTTCTCTGATTT 406
   ::::|
56 GlnTyrArgAlaCysProcysLeuArgAsnLeuThrCys 68
   ::::|
407 GAATCACAGCCATGCCCTGCAACATGAACACACAGTGC 445

seq_name: /cnrl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-970-966-206

seq_documentation_block:
; Sequence 206, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT FILING DATE: 2001-10-02
; CURRENT APPLICATION NUMBER: US/09/970,966
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 2364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-970-966-206

alignment_scores:
  Quality: 74.00      Length: 103
  Ratio: 1.321        Gaps: 4
  Percent Similarity: 54.369  Percent Identity: 20.388

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-970-966-206 ..
Align seg 1/1 to: US-09-970-966-206 from: 1 to: 2364

11 GluLeuLysGluSerCysTleArgAsnGlnAspCysGluThrGlyCysCy 27
   |||
1567 GAGGAACCGAAGTGCACGGTCAACGAGGAGTCTCTCCAGCAGCTG 1616
   |||

27 sGlnArgAlaPro.....AspAsnCysGluSerHisCysAla. 39
   |||
1751 TGGCATGAAGAAGCGCACCGCATGATCAAGATGAACCCCGCAGATGC 1800
   |||
46 SerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAlaCysProCy 62
   |||
1801 TCCATGTGCAAAAGCCGAGATCAGAGGAGAGTGCATGATGATGCCAGA 1850
   |||
62 sLeuArgAsnLeuThrCysTleTyrSerLysAsnGluLysTrp.....L 77
```



```
seq_documentation_block:
; Sequence 5020, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From cDNA Libraries Of Fetal Liver-Spleen and
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
```

Thu Nov 15 10:52:08 2001

us-09-215-435-179_copy_24_121.rnpn

```

; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5020
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-340-5020

```

```

alignment_scores:
  Quality: 71.50      Length: 84
  Ratio: 1.663       Gaps: 3
  Percent Similarity: 51.190  Percent Identity: 25.000

```

```

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-922-340-5020  ..

```

```

Align seg 1/1 to: US-09-922-340-5020 from: 1 to: 427

```

```

26 CysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysG1 42
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
44 TGCCTGCACATACCCAGTCCACTCTTTGATGCATTTAGTAAAGGGG 93
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
42 ySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgA 59
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
94 AGAGTCTGGGTGCTTTGCTACTGTTTCTATTCCGTAAGAAAAATGGCC 143
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
59 laCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 75
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
144 CATGAAGAAACTTAAAAACTCAAGTTTGAATTTCTCCAGGCAGGCAAAA 193
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
76 .....TrpLeuSerIle.....Al 80
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
194 ATGAGGTAATAAGTAAGGACTTGGCTGGCAATTCCTCAAGAGCAGG 243
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
80 aTyrglyArgCysGln...LysIleGlyArgGlnLysLeuAlaLysLysM 96
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
244 CTATGGAGGAGCAGGTCCTCCGAGTAGGTAGGAGGCTTATCGAGAGAAGC 293
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
96 et 96
::
294 TG 295

```

```

seq_name: /cgn1_7/ptodata/1/pna/us09_NEW_COMB.seq:US-09-922-340-5019

```

```

seq_documentation_block:
; Sequence 5019, Application US/09922340
; GENERAL INFORMATION:
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FROM CDNA LIBRARIES OF FETAL LIVER-SPLEEN AND INFANT BRAIN
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5019
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(461)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-5019

```

```

alignment_scores:
  Quality: 71.50      Length: 84
  Ratio: 1.663       Gaps: 3
  Percent Similarity: 51.190  Percent Identity: 25.000

```

```

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-922-340-5019  ..

```

```

Align seg 1/1 to: US-09-922-340-5019 from: 1 to: 461

```

```

26 CysCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysG1 42
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
44 TGCCTGCACATACCCAGTCCACTCTTTGATGCATTTAGTAAAGGGG 93
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
42 ySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgA 59
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
94 AGAGTCTGGGTGCTTTGCTACTGTTTCTATTCCGTAAGAAAAATGGCC 143
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
59 laCysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLys 75
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
144 CATGAAGAAACTTAAAAACTCAAGTTTGAATTTCTCCAGGCAGGCAAAA 193
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
76 .....TrpLeuSerIle.....Al 80
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
194 ATGAGGTAATAAGTAAGGACTTGGCTGGCAATTCCTCAAGAGCAGG 243
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
80 aTyrglyArgCysGln...LysIleGlyArgGlnLysLeuAlaLysLysM 96
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
244 CTATGGAGGAGCAGGTCCTCCGAGTAGGTAGGAGGCTTATCGAGAGAAGC 293
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
96 et 96
::
294 TG 295

```

```

seq_name: /cgn1_7/ptodata/1/pna/us60_NEW_COMB.seq:US-60-325-448-2381

```

```

seq_documentation_block:
; Sequence 2381, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 2381
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-2381

```

```

alignment_scores:
  Quality: 71.50      Length: 98
  Ratio: 1.521       Gaps: 5
  Percent Similarity: 47.959  Percent Identity: 26.531

```

```

alignment_block:
US-09-215-435-179_COPY_24_121 x US-60-325-448-2381/rev  ..

```

```

Align seg 1/1 to reverse of: US-60-325-448-2381 from: 1 to: 2000

```

```

6 LysTyrAsnLeuLeuGluLeuLysGluSerCysIleArgAsnGlnAspCY 22
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
1112 AAAAAGAATGTAGATACAAACAGAGCAAAACATCAGCAAGTCTACATG 1063
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
22 sGluThrGlyCysCysGlnArgAlaPro.....AspA 33
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
1062 TACAACCTGCTGCTATCAAAACTTCATTTATCTACAAGAACAAATTGATC 1013
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

```


5

```
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15

alignment_scores:
    Quality: 70.00      Length: 68
    Ratio: 2.059        Gaps: 3
    Percent Similarity: 50.000    Percent Identity: 26.471

alignment_block:
US-09-215-435-179_copy_24_121 x US-09-978-189-455 ..
Align seg 1/1 to: US-09-978-189-455 from: 1 to: 1518

16 CysIleArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAs 32
||||:||||:||||:||||:||||:||||:||||:||||:||||:
702 TGTCTCCGGTCATCAGACTGTGCTCAGGATTGTGT..... 737
32 pAsnCysGluSerHis.....CysAlaGluLysGlySerG 44
||||:||||:||||:||||:||||:||||:||||:||||:||||:
738 ....TGTGCTAGACACTTCTGCTCAAGATCTGTAACCTGCTCGAAG 783
44 lucIySerLeuCysGlnThrGlnValPheGlyGln..... 56
||||:||||:||||:||||:||||:||||:||||:||||:||||:
784 AAGGTCAAGTGTGTACCAAGCAGTAGAGAAAGGCTCTCATGGACTAGAA 833
57 ....TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerly 72
||||:||||:||||:||||:||||:||||:||||:||||:||||:
834 ATATTCCAGCGTTGTACTGTGGAGAAAGGTCTGTCTGCCGATACAGAA 883
72 sAsn 73
||||:
884 AGAT 887

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-192-455

seq_documentation_block:
; Sequence 455, Application US/09978192
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192
; CURRENT FILING DATE: 2001-10-15
```

```
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
```

```

PRIORITY APPLICATION NUMBER: 60/085338
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085323
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085582
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085700
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085689
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085579
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085580
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085573
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085704
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085697
PRIORITY FILING DATE: 1998-05-15

alignment_scores:
    Quality: 70.00
    Ratio: 2.059
    Percent Similarity: 50.000    Percent Identity: 26.471

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-978-192-455  ..
Align seg 1/1 to: US-09-978-192-455 from: 1 to: 1518

16 CysIleArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAs 32
   |||:||||:||||:||||:||||:||||:||||:||||:||||:
702 TGTCGCGGTCAACAGACTGTCCTCAGGATTGTG..... 737

32 pAsnCysGluSerHis.....CysAlaGluLysGlySerG 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738 ...TGTCGTAGACACTGTCGTCCAGATCTGTAAACCTGCTCTGAAAG 783

44 luGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 56
   |||:||||:||||:||||:||||:||||:||||:||||:
784 AAGCTCAAGTGTGTACCAAGCATAGGAGAAAAGGCTCTCATGGACTAGAA 833

57 ...TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerIy 72
   |||:||||:||||:||||:||||:||||:||||:||||:
834 ATATTCAGCGCTTGTTACTGTGGAGAGGCTGTGCTTCGCGATACAGAA 883

72 sasn 73
   |||:
884 AGAT 887

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-697-455

seq_documentation_block:
; Sequence 455, Application US/09978697
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.

```

APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 05/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637

us-09-215-435-179_copy_24_121.rnnp

Thu Nov 15 10:52:08 2001

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15

```

```

alignment_scores:
  Quality: 70.00      Length: 68
  Ratio: 2.059       Gaps: 3
  Percent Similarity: 50.000   Percent Identity: 26.471

```

```

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-978-697-455 ..
Align seg 1/1 to: US-09-978-697-455 from: 1 to: 1518

16 CysileArgAsnGlnAspCysGluThrGlyCysCysGlnArgAlaProAs 32
||||:||||:||||:||||:||||:||||:||||:||||:||||:
702 TGTCTCCGGTCATCAGACTGTGCTCCTCAGGATTGTG..... 737

32 pAsnCysGluSerHis.....CysAlaGluLysGlySerG 44
||||:||||:||||:||||:||||:||||:||||:||||:||||:
738 ....TGTGCTAGACACTTCTGTCTCAAGATCTGTAACCTCTCTGAAG 783

44 luGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 56
||||:||||:||||:||||:||||:||||:||||:||||:||||:
784 AAGGCAAGTGTGTACCAAGCATAGAGAAAGGCTCTCATGGACTAGAA 833

57 ...TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLy 72
||||:||||:||||:||||:||||:||||:||||:||||:||||:
834 ATATTCACCGTGTGTACTGTGTGAGAAAGGCTGTCTTCCCGGATACAGA 883

72 sAsn 73
||||:
884 AGAT 887

```

seq_name: /cgn1/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-824-455

```

seq_documentation_block:
; Sequence 455, Application US/09978824
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Av1
; APPLICANT: Baker Kevin P.

```

```

; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Olang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C14
; CURRENT APPLICATION NUMBER: US/09/978,824
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786

```



```
57 ...TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLy 72
      ::::: ||| ||| :::::|||||::: ||| ::: ||
834 ATATTCACAGCTGTGTACTGTGGAGAAGGTCTCTTCCCGGATACAGAA 883
      72 sAsn 73
      ||::
884 AGAT 887

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-974-298-88

seq_documentation_block:
; Sequence 88, Application US/09974298
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 88
; TYPE: DNA
; LENGTH: 1797
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 2481150CB1
US-09-974-298-88
```

```
alignment_scores:
Quality: 70.00 Length: 68
Ratio: 2.059 Gaps: 3
Percent Similarity: 50.000 Percent Identity: 26.471

alignment_block:
US-09-215-435-179_COPY_24_121 x US-09-974-298-88 ..
Align seg 1/1 to: US-09-974-298-88 from: 1 to: 1797

16 CysIleArgAsnGlnAspCysGluThrGlyCysGlnArgAlaProAs 32
   ||:::|||||:::|||||::: ||| |||
716 TGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTG..... 751
32 pAsnCysGluSerHis.....CysAlaGluLysGlySerG 44
   ||| |||
752 ....TGCTAGACACTTCTGCTCCAGATCTGTAACCTGCTCCTGAAG 797
44 LuGlySerLeuCysGlnThrGlnValPhePheGlyGln..... 56
   |||:::|||||::: |||
798 AAGGTCAGTGTGTACCAAGCATAGGAGAAAAGGCTCTCATGGACTAGAA 847
57 ...TyrArgAlaCysProCysLeuArgAsnLeuThrCysIleTyrSerLy 72
      ::::: ||| ||| :::::|||||::: ||| ::: ||
848 ATATTCACAGCTGTGTACTGTGGAGAAGGTCTCTTCCCGGATACAGAA 897
      72 sAsn 73
      ||::
898 AGAT 901
```



```

392 CCAGGACTGCGAGACTGCTGCTGCCACACGTCTCCAGACAAATGCGAGT 343
36 exHis.CysAlaGluLysGlySerGluLysSerLeuCysGlnThrGlnVa 52
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
342 CCAGCTNGCGCGGAGGAGGGTCCGAGCGAGCTGCTGTCAACACGAGGT 293
52 lPhePheGlyGlnTyrArgAlaCysProCysLeuArgAsnLeuThrCysI 69
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
292 GTTCCTTGGCCAAATATAGAGCGTGTCCCTGCTG.CGGAACCTGACTTGT 244
69 letYrSerLysAsnGluLysTyrPheSerIleAlaTyrGlyArgCysGln 85
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
243 TATATTCAAGAAGTGAAGTGGCTTAGCATGCTGCTATGGCGTGTGTGAC 194
86 LysIleGlyArgGlnLysLeuAlaLysLysMetPhePhe 98
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
193 AAAATTGGAGGAGGAGGTTGGCTAAGAAAATGTTCTTC 155

```

seq_name: gb_gss29:AZ575911

```

seq_documentation_block: 127 bp DNA GSS 06-DEC-2000
LOCUS AZ575911
DEFINITION AST-T22D0047 Genetrap T47D Human Breast Carcinoma Library Homo
sapiens genomic 5', DNA sequence.
ACCESSION AZ575911
VERSION AZ575911.1 GI:11562222
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 127)
AUTHORS Henkel,G., Liyanage,M., Pratt,E., Huang,D., Riley,M., Bernardino,A.,
Durick,K. and Pollok,B.
TITLE Exon-trap tags from a T47D GenomesScreen(TM) Library
JOURNAL Unpublished (2000)
COMMENT Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com
Pools of cells were isolated from a GenomesScreen(TM) library. The
library of cells was generated by retroviral integration of a gene
tagging element consisting of: 1) A promoterless beta-lactamase
proceeded by a splice acceptor as a reporter for gene expression;
2) A promoter driving neomycin resistance followed by a splice
donor to trap downstream exons. 3' RACE from neomycin gene was
performed using total RNA from isolated pools. Output was shotgun
cloned in pAMP-1 and used to transform DH5-alpha competent
bacteria. 5' ends of reported sequences were immediately preceded
by splice donor from the trapping construct.
Class: exon-trapped.

```

```

FEATURES
    source
        1..127
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="Genetrap T47D Human Breast Carcinoma Library"
            /tissue_type="Carcinoma"
            /cell_type="Epithelial"
            /cell_line="T47D"
            /note="Organ: Breast; Vector: pAMP-1; 3' RACE of total RNA
from genetrap pools; shotgun clone in pAMP-1 and used to
transform DH5-alpha competent bacteria."
BASE COUNT 28 a 34 c 43 g 21 t 1 others
ORIGIN

```

```

alignment_scores:
    Quality: 196.00 Length: 40
    Ratio: 5.026 Gaps: 0
    Percent Similarity: 97.500 Percent Identity: 97.500

```

```

alignment_block:
US-09-215-435-179_copy_24_121 x AZ575911
Align seg 1/1 to: AZ575911 from: 1 to: 127
11 GluLeuLysGlySerCysIleArgAsnGlnAspCysGluThrGlyCysCy 27
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 GAGCTCAGGAGTCTTGCTATCGGACACGAGGACTCGAGACTGCTGCTG 50
27 sGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlySerG 44
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 CCAACGTGCTCCAGACAA.TGCGAGTCCGCACTGCGCGGAGAGGGGTCCN 99
44 luGlySerLeuCysGlnThr 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
100 AGGCAGCTCTGTGTCAACG 119
seq_name: gb_est88:BF509739

```

```

seq_documentation_block: 293 bp mRNA EST 06-DEC-2000
LOCUS BF509739
DEFINITION UI-H-B14-agg-b-03-0-UI.s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone
IMAGE:3087296 3', mRNA sequence.
ACCESSION BF509739
VERSION BF509739.1 GI:11593037
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 293)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
The sequence contained an oligo-dr track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.

```

```

FEATURES
    source
        1..293
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NCI_CGAP_Sub8"
            /lab_host="DH10B (Life Technologies)"
            /note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; NCI_CGAP_Sub8
is a subtracted library derived from NCI_CGAP_Sub5. The
NCI_CGAP_Sub8 library had 2.5 million recombinants. A
single-stranded DNA preparation of NCI_CGAP_Sub5 was used
as a tracer in a subtractive hybridization with a driver
comprising: a pool of clones from NCI_CGAP_Sub5 (IMAGE
clone ids 2732833-2737415, 3068040-3069191, 25% of the
driver population), a pool of clones from NCI_CGAP_Sub4
(driver population), a pool of clones from NCI_CGAP_Sub5
(IMAGE clone ids 2723592-2729326, 25% of the driver
population), NCI_CGAP_Sub6 (pool AIF-AJU, IMAGE ids
2728969-2733190, 25% of the driver population), and
NCI_CGAP_Sub7 (IMAGE ids 3069192-3072238, 3081864-3084550
; 25% of the driver population). Subtraction was
performed as previously described [Bonaldi, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG_LIB=NCI_CGAP_Br2

```

```

TAG_TISSUE=breast
TAG_SEQ=AAACC
BASE COUNT      63 a      75 c      72 g      83 t
ORIGIN

alignment_scores:
  Quality: 164.50      Length: 38
  Ratio: 5.306        Gaps: 1
  Percent Similarity: 81.579      Percent Identity: 78.947

alignment_block:
US-09-215-435-179_COPY_24_121 x BF509739/rev ..
Align seg 1/1 to reverse of: BF509739 from: 1 to: 293

25 GlyCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLy 41
|||||
284 GCGTCTGCCACAGCTCCAGACAATTGCGAGTCGCACTGCGCGAGAA 235
|||||
41 sGlySerGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyra 58
|||||
234 GGGTCGAGGGCAGCTCTGTGTCACACGAGCTG.....C 200
|||||
58 rgAlaCysProCys 62
|||||
199 CAGCTACTCCATGC 186

seq_name: gb_est69:BE047055

seq_documentation_block:
LOCUS      BE047055      540 bp      mRNA      EST      08-JUN-2000
DEFINITION hg55g09.x1 NCI_CGAP_Pan3 Homo sapiens cDNA clone IMAGE:3123328 3'
            similar to gb:J02883 COLIPASE PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION  BE047055
VERSION    BE047055.1 GI:8364108
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 540)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 397.
            Location/Qualifiers
            1..540
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:3123328"
               /clone_lib="NCI_CGAP_Pan3"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.1 kb. Library constructed by Life
               Technologies."
BASE COUNT  123 a    131 c    161 g    124 t
ORIGIN

```

```

alignment_scores:
  Quality: 137.00      Length: 87
  Ratio: 2.362        Gaps: 2
  Percent Similarity: 66.667      Percent Identity: 32.184

alignment_block:
US-09-215-435-179_COPY_24_121 x BE047055/rev ..
Align seg 1/1 to reverse of: BE047055 from: 1 to: 540

10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrClyCy 26
|||||
437 TTGGAGACCGGTGAGCTGTCATGATAGTATGATGATGATGATGATGATG 388
|||||
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlys 43
|||||
387 CTGCCAGCATTCAGTGGCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 338
|||||
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyraGAla 59
|||||
337 GCGAGAACAGCGAGTCTCTGTCAAGACGCTTATGGGATTTACTACAAG 288
|||||
60 CysProCysLeuArgAsnLeuThrCysIleTySerLysAsnGluLysTr 76
|||||
287 TGTCCCTGTGAGCTGGCTGACCTGT.....GAGGGAGACAAGAC 247
|||||
76 pLeu.....SerIleAlaTyrglyArgCysGlnLysileG 88
|||||
246 CATTGTGGGCTCCATCACCACCACTTGGCATTTGCCATGACGNTG 197
|||||
88 lyArgGlnLys 91
|||||
196 GAGCTGCCAAG 186

seq_name: gb_est19:AI366744

seq_documentation_block:
LOCUS      AI366744      484 bp      mRNA      EST      13-FEB-1999
DEFINITION qq41a03.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1935052
            3' similar to gb:J02883 COLIPASE PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION  AI366744
VERSION    AI366744.1 GI:4136489
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 484)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 715      Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 445.
            Location/Qualifiers
            1..484
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:1935052"
               /clone_lib="Soares_NHMPu_S1"
               /tissue_type="Pooled human melanocyte, fetal heart, and
               pregnant uterus"
               /lab_host="DH10B"
               /note="Organ: mixed (see below); Vector: pT7T3D-Pac
               (Pharmacia) with a modified polylinker; Site_1: Not I;
               Site_2: Eco RI; Equal amounts of plasmid DNA from three
               normalized libraries (melanocyte 2NbHM, pregnant uterus
               NbHPU, and fetal heart NbHH19W) were mixed, and ss circles

```

BASE COUNT
ORIGIN

```

alignment_scores:      Length:      87
                       Quality: 135.00      Gaps:      2
                       Ratio:    2.328
                       Percent Similarity: 66.667      Percent Identity: 32.184

alignment_block:
000355435-179_CBPV_24_121_x_AI366744/rev ..

```

Alian seg 1/1 to reverse of: AI366744 from: 1 to: 484

10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||||::: :|||:::|
23 CCCCCCAAGCCCTCAGCTCTGCATGAATAGTGTGCCAGTGTAAGAGCAATTG 374

26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlyS 43
|||||::: ::::|::: :::
::: ::: ::: ::: :::

43•erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyArgAla 59

60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
||||| ||||| :|||
:|::|:|:|:

273 TGTCCTGTGAGCGTGGCCTGACCTGT.....GAGGGAGAGCAGCAGC
76 pLeu.....SerIleAlaTyRGlyArgCysGlnLysIleG 88

232 CATCGTGGGCTCCATCACCAACCAACTTTGGCATCTGGCCATGACGCTG 18
88 lyArgGlnLys 91

182 GACGCTCCAAG 172

seq_documentation_block:					
LOCUS	BF507781	499 bp	mRNA	EST	06-DEC-2000
DEFINITION	UI-H-BI4-apv-b-04-o-ui.sl NCI_CGAP_Sub8 Homo sapiens cDNA clone				
FEATURES	unlabeled primer location 1..3' mRNA sequence.				

ACCESSION	BF507781	
VERSION	BF507781.1	GI:11591079

KEYWORDS EST.
SOURCE human.

ORGANISM	Homo Sapiens
	Eukaryote

REFERENCE : 1 (ba

AUTHORS	NATIONALITY	TITLE
...

JOURNAL

Email

oligon
strantail.
NCI-C

L.M.A
www-b

seq. p
POLYA

1

FEATURES	Location/Qualifiers
source	1. .499

```

/clonelib="NCI_CGA
/clonelib="IMAGE:30884
/db_xref="taxon:960
/clonelib="NCI_CGA

```

```

/rab_host= dnv02 (
/note="Vector: pT7T
polylinker; Site_1:
is a subtracted lib

```

NCI_CGAP_Sub8 library
single-stranded DNA
as a tracer in a sub-

```
clone Ids` 2732833-2
driver population),
(IMAGE clone Ids 27
```

NCI_CGAP_Sub7 (IMAG
population); NCI_

performed as previously
Soares (1996): Norm
Approaches To Facilitate
1997, 1998

TAG_LIB=NCI_CGAP_BF
TAG_TISSUE=brain
TAG_SEQ=ATAGG"

BASE COUNT	106 a	121 c	133 b
ORIGIN			

```
alignment_scores:      135.00
Quality:              2.328
Ratio:
```

```
alignment_block:
US-09-215-435-179_COPY_24_121 x BF50777
```

Align seg 1/1 to reverse of: BF507781

439 CTGGAGAACGGTGAGCTCTGCATGAATAGT

389 CTGCCAGCATTCAAGTCCGCTGGGCCTGGCG

43 **ergtuglyscilccg**
|||||:|||| |:::~::~:
339 GCGAGAACAGCGAGTGCTGTGTC AAGACGGC

60 CysProCysLeuArgAsnLeuIrnCysIle
 ||||| |||::|||
 289 TGTCCCTGTGAGCGTGGCCTGACCTGT...

76 pLeu.....SerIleAlaT
:::
248 CATCGTGGGCTCCATCACCAACACCAACT

88 lyArgGlnLys 91
 ||||:||||
 108 CACGCTCCAAG 188

```
seq_name: gb_est81:BE969660
```

LOCUS	BE969660	528 bp	mr
DEFINITION	601679794F1	NIH_MGC_78	Hom
mRNA sequence.			

ACCESSION BE303000

1

88 lyArgGlnLys 91

```

277 TGTCCCTGTGACGCTGGCTGACCTGT.....GAGGAGACAAGAC 237
76 pLeu.....SerileAlaTyrGlyArgCysGlnLysIleG 88
236 CATCGTGGGCTCCATCAGCAACCAACTTTGGCATCTGCCATGACGCTG 187
88 lyArgGlnLys 91
186 GAGGCTCCAAG 176

seq_name: gb_est81:BE969913
seq_documentation_block: 533 bp mRNA EST 04-OCT-2000
LOCUS BE969913
DEFINITION 601680001F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950113 5',
mRNA sequence.
ACCESSION BE969913
VERSION BE969913.1 GI:10582846
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 533)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM816 row: b column: 02
High quality sequence stop: 487.
FEATURES
Location/Qualifiers
1..533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950113"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGGCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 118 a 159 c 132 g 114 t
ORIGIN

alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184

alignment_block:
US-09-215-435-179_copy_24_121 x BE969913 ..
Align seg 1/1 to: BE969913 from: 1 to: 533
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyC 26
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 CTGGAGACGGTGAGCTCTGTCATGATATGTCGCCAGGAGCAATG 147

```

```

26 scysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlys 43
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 CTCCAGCATTTAAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 197
43 erGluGlySerLeuCysGlnThrGlnValPheGlyGlnTyrArgAla 59
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 GCGAGACAGCGAGTCTCTGTCAAGACGCTCTATGGGATTTACTACAAG 247
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 TGTCCCTGTGAGGCTGGCTGACCTGT.....GAGGAGACAAGAC 288
76 pLeu.....SerileAlaTyrGlyArgCysGlnLysIleG 88
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 CATCGTGGGCTCCATCAGCAACCAACTTTGGCATCTGCCATGACGCTG 338
88 lyArgGlnLys 91
339 GAGGCTCCAAG 349

seq_name: gb_est81:BE969784
seq_documentation_block: 543 bp mRNA EST 04-OCT-2000
LOCUS BE969784
DEFINITION 601679544F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3949859 5',
mRNA sequence.
ACCESSION BE969784
VERSION BE969784.1 GI:10582717
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM815 row: g column: 12
High quality sequence stop: 477.
FEATURES
Location/Qualifiers
1..543
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949859"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGGCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 118 a 176 c 133 g 115 t
ORIGIN

alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184

```


full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 139 a 174 c 144 g 120 t
ORIGIN

alignment_scores:
Quality: 135.00 Length: 87
Ratio: 2.328 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 32.184

alignment_block:
US-09-215-435-179_COPY_24_121 x BE969920 ..

Align seg 1/1 to: BE969920 from: 1 to: 577

```
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
98 CTGGAGACGGTGAGCTCTGCATGATAGTGCCTGTAAGAGCAATTG 147
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlys 43
||||| |||:||||:||||:||||:||||:||||:||||:||||:
148 CTGCCAGCATTCAGTGCCTGGCGCTGGCCCTGCACATCCATCGGCCA 197
||||| |||:||||:||||:||||:||||:||||:||||:||||:
43 exGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
198 GCGAGAACAGCGAGTCTCTGTCAAGACGCTCTATGGGATTTACTACAAG 247
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
248 TGTCCCTGTGAGCGTGGCTGACCTGT.....GAGGGAGACAAGAC 288

76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 88
289 CATCTGGGGCTCATCACCAACCACTTTGGCATCTGCCATCTGCCATGACGCTG 338

88 lyArgGlnLys 91
||||| |||:||||:
339 GAGGCTCCAAG 349
```

seq_name: gb_est81:BE969726

seq_documentation_block:
LOCUS BE969726 682 bp mRNA EST 04-OCT-2000
DEFINITION 601679676F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3949891 5',
mRNA sequence.

ACCESSION BE969726
VERSION BE969726.1 GI:10582659
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM815 row: h column: 20
High quality sequence stop: 547.
Location/Qualifiers
1. 577

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949891"

alignment_block:

US-09-215-435-179_COPY_24_121 x BE969784 ..

Align seg 1/1 to: BE969784 from: 1 to: 543

```
10 LeuGluLeuLysGluSerCysIleArgAsnGlnAspCysGluThrGlyCy 26
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
98 CTGGAGACGGTGAGCTCTGCATGATAGTGCCTGTAAGAGCAATTG 147
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlys 43
||||| |||:||||:||||:||||:||||:||||:||||:||||:
148 CTGCCAGCATTCAGTGCCTGGCGCTGGCCCTGCACATCCATCGGCCA 197
||||| |||:||||:||||:||||:||||:||||:||||:||||:
43 exGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyrArgAla 59
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
198 GCGAGAACAGCGAGTCTCTGTCAAGACGCTCTATGGGATTTACTACAAG 247
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
||||| ||| |||:||||:||||:||||:||||:||||:||||:||||:
248 TGTCCCTGTGAGCGTGGCTGACCTGT.....GAGGGAGACAAGAC 288

76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleG 88
289 CATCTGGGGCTCATCACCAACCACTTTGGCATCTGCCATCTGCCATGACGCTG 338

88 lyArgGlnLys 91
||||| |||:||||:
339 GAGGCTCCAAG 349
```

seq_name: gb_est81:BE969920

seq_documentation_block:
LOCUS BE969920 577 bp mRNA EST 04-OCT-2000
DEFINITION 601680009F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950115 5',
mRNA sequence.

ACCESSION BE969920
VERSION BE969920.1 GI:10582853
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM816 row: b column: 04
High quality sequence stop: 534.
Location/Qualifiers
1. 577

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3950115"
/clone_lib="NIH_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggccattatgcc); Site_2: SfII (ggccattatgcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCCGGCGGCGACATG-3' (30)BN-3'
(where B = A, C, G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for

[illegible]

```

LOCUS       BE970031                752 bp      mRNA              04-OCT-2000
DEFINITION   601680144F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950266 5',
             mRNA sequence.
ACCESSION    BE970031
VERSION      BE970031.1 GI:10582964
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 752)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      NIH-MGC http://mgs.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished (1999)
             Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-re@mail.nih.gov
             Tissue Procurement: CLONETECH Laboratories, Inc.
             CDNA Library Preparation: CLONETECH Laboratories, Inc.
             CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM816 row: a column: 11
             High quality sequence stop: 499.
FEATURES             Location/Qualifiers
             source          1..733
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone_image="IMAGE:3950098"
                 /clone_lib="NIH_MGC_78"
                 /lab_host="DH10B (T1 phage-resistant)"
                 /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
                 Site.1: SfII (ggcgcctcgcc); Site.2: SfiI (ggccattatgccc
                 ); 5' and 3' adaptors were used in cloning as follows: 5'
                 adaptor sequence: 5'-CAGGCCCATTTGGCC-3' and 3' adaptor
                 sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dt(30)BN-3'
                 (where B = A, C, or G and N = A, C, G, or T). Average
                 insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
                 contained inserts by PCR. This library was enriched for
                 full-length clones and was constructed by Clontech
                 Laboratories (Palo Alto, CA)."
BASE COUNT    171 a 218 c 210 g 134 t
ORIGIN
alignment_scores:
    Quality: 135.00      Length: 87
    Ratio: 2.328         Gaps: 2
    Percent Similarity: 66.667      Percent Identity: 32.184
alignment_block:
US-09-215-435-179_COPY_24_121 x BE970098 ..
Align seg 1/1 to: BE970098 from: 1 to: 733
10 LeuGlLeuLysGluSerCysIleArgAsnGlnAsnCysGluThrGlyCy 26
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 CTGGAGAACGGTGAGCTGTGCATGAATAGTGCCTCAAGACAAATG 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
26 sCysGlnArgAlaProAspAsnCysGluSerHisCysAlaGluLysGlys 43
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 CTGCCACATTCAGTGCCTGCGCTGGCCCTGCGCCCTGCATCCATCCAGCA 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 erGluGlySerLeuCysGlnThrGlnValPhePheGlyGlnTyArgAla 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 GCGAAGAACAGCGTGTCTGTCAAGACGCTCTATGGATTACTACAAG 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
60 CysProCysLeuArgAsnLeuThrCysIleTyrSerLysAsnGluLysTr 76
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 TGTCCTGTGAGCGTGGCTGTACCTGT .....GAGGAGACAAAGAC 288
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
76 pLeu.....SerIleAlaTyrGlyArgCysGlnLysIleg 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
289 CATCGTGGCTCCATCACCAACACCACTTTGCCATCTGCCATGAGCGTG 338
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 lyArgGlnLys 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 GAGCTCCAAG 349
seq_name: gb_est81.BE970031
seq_documentation_block:

```

```

:::
289 CATCGTGGGCTCCATCACCACCACTTTGGCATCTGCCATGACGCTG 338
88 lyArgGlnLys 91
   |||||:::|
339 GACGCTCCAG 349
```

OM of: US-09-215-435-225 to: GenEmbl: * out_format : pfs
Date: Nov 15, 2001 12:48 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framed_p2n.model -DRV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435_@CGNI_1_11006 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-215-435-225
Query length: 227
Database: GenEmbl: *
Database sequences: 1344157
Database length: -856060004
Search time (sec): 7698.940000

score_list:	Strd Orig	Zscore	EScore	Len	Documentation
Sequence					
gb_pat2:AX083412	+ 1224.00	2255.32	2.2e-117	681	! AX083412 Sequence 104 from Pat
gb_pat2:AX083422	+ 1224.00	2252.82	3.0e-117	897	! AX083422 Sequence 114 from Pat
gb_pat1:AX060293	+ 1220.00	2245.07	8.2e-117	932	! AX060293 Sequence 11 from Pat
gb_sts2:G27363	- 444.00	816.59	3.0e-37	352	! G27363 human STS SHGC-31033, se
gb_btgt13:AC055854	- 262.50	424.04	2.2e-15	179814	! AC055854 Homo sapiens chromo
gb_btgt12:AC037441	- 246.00	394.38	9.9e-14	162795	! AC037441 Homo sapiens clone
gb_btgt17:AC084081	- 246.00	394.01	1.0e-13	169645	! AC084081 Homo sapiens chrom
gb_to2:RNPBP	+ 237.50	424.35	2.1e-15	1037	! X71873 R.norvegicus mRNA for R
gb_pat2:E05646	+ 237.50	424.26	2.1e-15	1047	! E05646 cDNA encoding rat hippo
gb_to2:RNPABP	+ 237.50	424.02	2.2e-15	1075	! X75253 R.norvegicus phosphatid
gb_in4:OVDLP1PROTN	+ 234.50	423.57	2.3e-15	611	! X87991 O.volvulus mRNA for di
gb_in3:CELF4A03	+ 230.00	377.18	9.0e-13	41155	! AF016423 Caenorhabditis eleg
gb_btgt1:AC006915	+ 230.00	366.50	3.5e-12	134095	! AC006915 Caenorhabditis eleg
gb_btgt4:AC012978	- 229.00	371.40	1.9e-12	63551	! AC012978 Drosophila melanoga
gb_in1:AC005891	+ 229.00	364.46	4.6e-12	136933	! AC005891 Drosophila melanoga
gb_in1:AF003637	+ 229.00	360.02	8.1e-12	223646	! AE003637 Drosophila melanoga
gb_to1:AF226629	+ 227.50	404.47	2.7e-14	1205	! AF226629 Rattus norvegicus pho
gb_to1:AF300422	+ 225.00	406.70	2.0e-14	564	! AF300422 Mus musculus phosphati
gb_to1:AB046417	+ 225.00	399.86	4.9e-14	1202	! AB046417 Mus musculus HCNPpp n
gb_btgt6:AC020316	- 222.50	356.99	1.2e-11	82631	! AC020316 Drosophila melanoga
gb_in4:AE003675	+ 222.50	351.85	2.3e-11	145786	! AE003675 Drosophila melanoga
gb_in4:OVD2PROTN	+ 219.00	396.43	7.6e-14	514	! X87990 O.volvulus mRNA for di
gb_to1:MMU43206	+ 219.00	388.92	2.0e-13	1180	! U43206 Mus musculus phosphatid
gb_pr10:HMWRBPB	+ 208.50	367.71	3.0e-12	1434	! D16111 Human mRNA for human bc
gb_pr9:HSRBPB	+ 208.50	367.64	3.1e-12	1444	! X75252 H.sapiens phosphatidyle
gb_pr10:MFEBP	+ 208.50	367.64	3.1e-12	1444	! X73137 M.fascicularis mRNA for
gb_pat2:E05647	+ 208.50	367.62	3.1e-12	1447	! E05647 cDNA encoding human hip
em_pat:R09006	+ 207.50	369.10	2.5e-12	1001	! E09006 Human cDNA encoding a p
gb_btgt5:AC016445	+ 207.00	377.25	5.4e-10	92569	! AC016445 Drosophila melanoga
gb_pat2:AF332406	+ 206.50	373.14	1.5e-12	522	! AF332406 Arabidopsis thaliana p
gb_pat2:AF147721	+ 206.50	368.96	2.6e-12	829	! AF147721 Arabidopsis thaliana p
em_pat:E09005	+ 205.00	369.66	2.4e-12	564	! E09005 Human cDNA encoding a ph
gb_pr9:HSRNPBP	+ 205.00	369.56	2.4e-12	570	! X85033 H.sapiens mRNA for phosph
gb_pr10:SF6773	+ 202.00	361.86	6.4e-12	723	! SF6773 neuropolyptide h3 [hum
gb_pat2:111816	+ 201.50	359.77	8.4e-12	822	! I11816 Sequence 1 from Patent t
gb_in4:ONGANTIG	+ 201.50	359.77	8.4e-12	822	! M27807 O.volvulus O.V.-16 anti
gb_in1:AB035321	+ 196.00	353.48	1.9e-11	534	! AB035321 Dirofilaria immitis m
gb_btgt13:AC062037	- 194.00	296.80	2.7e-08	187355	! AC062037 Homo sapiens chromo
gb_btgt12:AC037441	- 191.50	293.44	4.2e-08	162795	! AC037441 Homo sapiens clone
gb_btgt5:AC016445	- 190.50	296.69	2.7e-08	92569	! AC016445 Drosophila melanoga
gb_pat2:AF152907	+ 182.00	327.66	5.2e-10	528	! AF152907 Arabidopsis thaliana t

gb_pl1:AB027506 + 182.00 323.91 8.3e-10 799 ! AB027506 Arabidopsis thalian
gb_btgt4:AC013131 + 182.00 291.03 5.7e-08 30334 ! AC013131 Drosophila melano
gb_in1:AC004439 + 182.00 281.82 1.8e-07 84055 ! AC004439 Drosophila melano
gb_in1:AE003586 + 182.00 270.13 8.3e-07 306352 ! AE003586 Drosophila melan

seq_name: gb_pat2:AX083412

seq_documentation_block:
LOCUS AX083412 681 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 104 from Patent WO0112660.
ACCESSION AX083412
VERSION AX083412.1 GI:13185249
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Kato,S. and Kimura,T.
TITLE Human proteins having hydrophobic domains and dnas encoding these
proteins
JOURNAL Patent: WO 0112660-A 104 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
FEATURES
Location/Qualifiers
source 1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 179 a 186 c 183 g 133 t
ORIGIN

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x AX083412 ..

Align seg 1/1 to: AX083412 from: 1 to: 681

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuLeuLeuLeuMe 17
|||||
1 ATGGGTGGACAATGAGCTGGTGCACAGCAGCAGCTCTTACTGGGTCTCAT 50
17 MetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
51 GATGTTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 100
34 laLeuLeuAspGluAspThrLeuPheCysGlnGluValPheTyr 50
|||||
101 CCTCTTTGACGAGGACACCCCTTTTTCAGGCGCTTGAAGTTTCTAC 150
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnty 67
|||||
151 CCAGAGTTGGGGAACATTGGCTGCAAGGTTGTCTCGATTGTAACAACTA 200
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
201 CAGACAGAGATCACCTCTCGTGGATGGAGCGATGATCAAGTTCCTCCGGGG 250
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProaspAlaPro 100
251 CCGTGGACGGCGCAACCTATATCTCTGGTGGATGGATCCAGATGCCCT 300
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
301 ACCAGACGACAGACCCACAGACAGAGATTCTGGAGACATTGGCTGTAACACA 350
117 pLeLysGlyAlaAspLeuLysGlyLysLeGlnGlnGlnLeuLeus 134
351 TATCAAGGGCGCGACCTCAAGAAGGAAGATTTCAGGCCAGGAGTTAT 400

Thu Nov 15 10:52:08 2001

us-09-215-435-225.rge

134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
401 CAGCCTACAGGCTCCCTCCACCAGCGGACACAGTGCTTCATCGCTAC 450
151 GlnPhePheValTyrLeuGlnGluGlyLysValLleSerLeuLeuProLy 167
451 CAGTCTTTGTCTATCTTCAGAGGAAAGTTCATCTCTCTCTCTCCAA 500
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
501 GGAAACAAACACTCGAGGCTCTTGGAATATGGACAGATTTCTGAACCGTT 550
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
551 TCCACCTGGCGCAACTGAAGCAAGCACCAGTTCATGACCCAGAACTAC 600
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
601 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAA 650
217 sHisLysAsnGlnAlaGluLeuAlaAlaCys 227
651 GCACAAAACCCAGCGGAGATAGTGCCTGC 681

seq_name: gb_pat2:AX083422

seq_documentation_block: 897 bp DNA PAT 28-FEB-2001
LOCUS AX083422
DEFINITION Sequence 114 from Patent WO0112660.
ACCESSION AX083422
VERSION AX083422.1 GI:13185262
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS Kato,S. and Kimura,T.
TITLE Human proteins having hydrophobic domains and dnas encoding these proteins
JOURNAL Patent: WO 0112660-A 114 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP); Protegene Inc. (JP)
FEATURES
source
1..897
/organism="Homo sapiens"
/db_xref="taxon:9606"
99..782
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC33305.1"
/db_xref="GI:13185263"
/translation="MGWTWRLVTAALLLGLMMVTVGDEDENSPCAHEALLDLDLFCQ
GLVFFYFELGNIGCKVPPDCNNYRKTITSMMEPIVKFPGAVDGTATILVMVDPDAPSR
AEPQRFWRHLWDIKGADLKKGITQGOELSAOAPSPAHSGFHRQFFVYLQEGK
VISLLEPKNTKRWGMDKRLNRLPHLGEPEASTQFNTQYQDSPTLQAPRERASEPKH
KNQAEATAAC"

BASE COUNT 226 a 253 c 234 g 184 t
ORIGIN
alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225 x AX083422 ..
Align seg 1/1 to: AX083422 from: 1 to: 897
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
99 ATGGGTTGGACAATGAGGCTGGTGCACAGCAGCACTGTTACTGGGTCTCAT 148

17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
149 GATGGTGGTCTACTGGAGACGAGATGAGAACAGCCCGTGTGCCCATGAGG 198
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
199 CCCTCTTGGACGAGGACACCCCTTTTGCAGGGCCTTGAAGTTTCTTCTAC 248
51 ProGluLeuGlyAsnLleGlyCysLysValValProAspCysAsnAsnTly 67
249 CCAGAGTTGGGGAACATTTGGCTGCAAGGTTGTTCCGATTTGTAACAAC 298
67 rArgGlnLysIleThrSerTrpMetGluProLleValLysPheProGlyA 84
299 CAGACAGAAGATCACCTCTCTGGATGGAGCGGATAGTCAAGTTTCCCGGGG 348
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
349 CCGTGGACGGCGCAACCTATATCTGTGATGTGGATCCAGATGCCCT 398
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
399 AGCAGACAGAACCCACAGACAGATTTCTGGAGACATTTGGCTGGTAACAG 448
117 pIleLysGlyAlaAspLeuLysLysGlyLysLleGlnGlyGlnGluLeuS 134
449 TATCAAGGGCGCCACCTGAAGAAAGGAAAGTTCAGGGCCAGGAGTTAT 498
134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
499 CAGCCTACCGGCTCCCTCCCGCCGACACAGTGGCTTCCATCGCTAC 548
151 GlnPhePheValTyrLeuGlnGlyLysValLleSerLeuLeuProLy 167
549 CAGTCTTTGTCTATCTTCAGAGAAAGAAAGTCTCTCTCTCTTCCCAA 598
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
599 GGAAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTTCTGAACCGTT 648
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
649 TCCACCTGGCGCAACCTGAAGCAAGCACCAGTTCATGACCCAGAACTAC 698
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
699 CAGGACTCACCAACCCCTCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAA 748
217 sHisLysAsnGlnAlaGluLeuAlaAlaCys 227
749 GCACAAAACCCAGCGGAGATAGTGCCTGC 779
seq_name: gb_pat1:AX060293

seq_documentation_block: 932 bp DNA PAT 22-JAN-2001
LOCUS AX060293
DEFINITION Sequence 1 from Patent WO0078802.
ACCESSION AX060293
VERSION AX060293.1 GI:12405782
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS Shimkets,R.A., Fernandes,E., Vernet,C., Yang,M., Boldog,F.L. and
Herrmann,J.L.
TITLE Secreted polypeptides and corresponding polynucleotides
JOURNAL Patent: WO 0078802-A 1 28-DEC-2000;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..932

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
113..796
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC24866.1"
/db_xref="GI:12405783"
/translation="MGWMTMLVTAALLGLMVVTGDEENSPCAHEALDEDTLFCQ
GLEVFPELGNICKVVPDCNNYRKITSWMEPIVKFPGAVDGTATILNMVDPDAPSR
AEPRORFWRHLEWTDLIKADLKEGKIQGOELSAVQAPSPAHSGFHYOFVYLOEGK
VISLLPKENKTRGSWMKDRFLNRFHLGEPASTQFMQNTQYQDSFTLQAPRERASEPKH
KNQAEIAAC"
244 a 257 c 244 g 187 t
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 1220.00 Length: 227
  Ratio: 5.374 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.559

alignment_block:
US-09-215-435-225 x AX060293 ..
Align seg 1/1 to: AX060293 from: 1 to: 932
1 MetGlyTTPThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
113 ATGGGTGGACAAATGAGGCTGTCACAGCAGCAGCTTACTGGGCTCAT 162
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
163 GATGTGTGTCTACTGGAGCAGAGATGAGAACAGCCGCTGTGCCCATGAGG 212
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
213 CCCTCTTGACGAGGACACCCCTCTTTGCCAGGGCCTTGAAGTTTCTAC 262
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnTyr 67
263 CCAGAGTTGGGGAACATGGCTGCAAGGTTGTCTCTGATTGTAACAATA 312
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
313 CAGACAGAAGATCACCTCTCTGATGGAGCGGATAGTCAAGTTCCTGGGG 362
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
363 CCGTGGAGCGGCGCAACCTATATCTCTGGTGTGATGATCCAGATGCCCT 412
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
413 ACACAGCAGCAACCCACAGACAGATTCGGACACATTTGGCTGGTAACAGA 462
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeus 134
463 TATCAAGGCGCCGACCTGAAGAGGAAGATTTCAGGSCCAGGAGTTAT 512
134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
513 CAGCCTACCAAGCTCCCTCCCGCCAGCGGACACAGTGGCTTCCATCGCTAC 562
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProly 167
563 CAGTCTTGTCTATCTTCAGGAAGGAAGATCATCTCTCTCTCTCCCAAA 612
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
613 GGAAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTCTGAACCGTT 662
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
663 TCCACCTGGGCGAACCTGAAGCAAGCACCACCTTCATGATCACCAGAACTAC 712
```

```
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
|||||
713 CAGGACTCACCACCTCCAGGCTCCAGAGAAAGGCCGAGGAGCCCAA 762
```

```
217 shisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
763 GCACAAAACACGAGCGGAGATAGCTGCCTGC 793
```

seq_name: gb_sts2:G27363

seq_documentation_block:
LOCUS G27363 352 bp DNA STS
DEFINITION human STS SHGC-31033, sequence tagged site.
ACCESSION G27363
VERSION G27363.1 GI:1396086
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Myers,R.M.
JOURNAL Unpublished (1996)
COMMENT

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: TTGGAAGAAAGGGTCTC
Primer B: AGCACAAAACCCGCG
STS size: 133
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds
Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H83927
-- Washington University/Merck EST sequence.
Location/Qualifiers
1. 352
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"
4. 136
primer_bind
primer_bind complement(119..136) 96 t 1 others
BASE COUNT 69 a 76 c 110 g 96 t 1 others
ORIGIN

alignment_scores:
Quality: 444.00 Length: 83

Thu Nov 15 10:52:08 2001

Ratio: 5.349 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x G27363/rev ..

Align seg 1/1 to reverse of: G27363 from: 1 to: 352

145 SerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGluGlyLysVa 161
|||||
352 AGTGCGTTCATCGCTACCGATCTTCTCTATCTCAGGAAGGAAAGT 303
|||||
161 lIleSerLeuLeuProLysGluAsnLysThrArgGlySerTrpLysMeta 178
|||||
302 CAPCTCTCTCTCTCCCAAGGAACAAACTCGAGCTCTTGGAATGG 253
|||||
178 spArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerThrGln 194
|||||
252 ACAGATTCTGACCGTTTCCACCTGGCGGAACCTGAAGCAAGCACCCAG 203
|||||
195 PheMetThrGlnAsnTyrGlnAspSerProThrLeuGlnAlaProArgG1 211
|||||
202 TTCATGACCCAGAACCTACCGAGGACTCACCAACCTCCAGGCTCCAGAGA 153
|||||
211 uArgAlaSerGluProLysHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
152 AAGGCCAGCGAGCCCAAGCACAAAAACCAAGCGGAGATAGCTGCCTGC 104
|||||

seq_name: gb_htgl3:AC055854

seq_documentation_block:
LOCUS AC055854 179814 bp DNA HTG 25-MAY-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-459E5 map 8, WORKING DRAFT
SEQUENCE, 26 unordered pieces.
AC055854
VERSION AC055854.3 GI:8084680
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 179814)
Homo sapiens chromosome 8, Clone RP11-459E5
Unpublished
2 (bases 1 to 179814)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Callymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferrelira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,W.
Direct Submission
Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7770648.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9765
Center clone name: 459_E_5
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 167541 bases at least Q40
Consensus quality: 173729 bases at least Q30
Consensus quality: 175997 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 177314; sum-of-contigs
Quality coverage: 4.2 in Q20 bases; agarose-fp
Quality coverage: 4.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1730: contig of 1730 bp in length
* 1731 1830: gap of 100 bp
* 1831 3230: contig of 1400 bp in length
* 3231 3330: gap of 100 bp
* 3331 5675: contig of 2345 bp in length
* 5676 5775: gap of 100 bp
* 5776 7874: contig of 2099 bp in length
* 7875 7974: gap of 100 bp
* 7975 10477: contig of 2503 bp in length
* 10478 10577: gap of 100 bp
* 10578 12553: contig of 1976 bp in length
* 12554 12653: gap of 100 bp
* 12654 16386: contig of 3733 bp in length
* 16387 16486: gap of 100 bp
* 16487 19715: contig of 3229 bp in length
* 19716 19815: gap of 100 bp
* 19816 23166: contig of 3351 bp in length
* 23167 23266: gap of 100 bp
* 23267 26855: contig of 3589 bp in length
* 26856 26955: gap of 100 bp
* 26956 31690: contig of 4735 bp in length
* 31691 31790: gap of 100 bp
* 31791 38834: contig of 7044 bp in length
* 38835 38934: gap of 100 bp
* 38935 46670: contig of 7736 bp in length
* 46671 46770: gap of 100 bp
* 46771 52635: contig of 5865 bp in length
* 52636 52735: gap of 100 bp
* 52736 59853: contig of 7118 bp in length
* 59854 59953: gap of 100 bp
* 59954 66887: contig of 6934 bp in length
* 66888 66987: gap of 100 bp
* 66988 75106: contig of 8119 bp in length
* 75107 75206: gap of 100 bp
* 75207 84076: contig of 8870 bp in length
* 84077 84176: gap of 100 bp
* 84177 91487: contig of 7311 bp in length
* 91488 91587: gap of 100 bp
* 91588 98734: contig of 7147 bp in length
* 98735 98834: gap of 100 bp
* 98835 107787: contig of 8953 bp in length
* 107788 107887: gap of 100 bp
* 107888 118191: contig of 10304 bp in length
* 118192 118291: gap of 100 bp

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liou,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meljirim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (08-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:8705171.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 19203
Center clone name: 87_E_22
----- Summary Statistics
Sequencing vector: M13; M7815; 81% of reads
Chemistry: Dye-terminator Big Dye; 19% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157025 bases at least Q40
Consensus quality: 160134 bases at least Q30
Consensus quality: 161093 bases at least Q20
Insert size: 158000; agarose-fp
Insert size: 161795; sum-of-contigs
Quality coverage: 5.4 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 ba.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 15243: contig of 15243 bp in length
* 15244 15343: gap of 100 bp
* 15344 18897: contig of 3554 bp in length
* 18898 18997: gap of 100 bp
* 18998 25211: contig of 6214 bp in length
* 25212 25311: gap of 100 bp
* 25312 34838: contig of 9527 bp in length
* 34839 34938: gap of 100 bp
* 34939 45351: contig of 10413 bp in length
* 45352 45451: gap of 100 bp
* 45452 57259: contig of 11808 bp in length
* 57260 57359: gap of 100 bp
* 57360 67990: contig of 10631 bp in length
* 67991 68090: gap of 100 bp
* 68091 83398: contig of 15309 bp in length
* 83400 83499: gap of 100 bp
* 83500 105395: contig of 21896 bp in length
* 105396 105495: gap of 100 bp
* 105496 129959: contig of 24464 bp in length
* 129960 130059: gap of 100 bp
* 130060 162795: contig of 32736 bp in length.

FEATURES
source
1..162795
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="RP11-87E22"
/clone_lib="RPC1-11 Human Male BAC"
1..15243
/note="assembly_fragment"
vector_end:SP6
vector_side:left
15344..18897
/note="assembly_fragment"
18898..18997
/note="assembly_fragment"
18998..25211
/note="assembly_fragment"
25212..34838
/note="assembly_fragment"
34839..45351
/note="assembly_fragment"
45352..57259
/note="assembly_fragment"
57260..67990
/note="assembly_fragment"
67991..83399
/note="assembly_fragment"
83400..105395
/note="assembly_fragment"
105396..129959
/note="assembly_fragment"
129960..162795
/note="assembly_fragment"
clone_end:T7
vector_side:right

BASE COUNT 42087 a 39347 c 38304 g 42057 t 1000 others
ORIGIN

alignment_scores:
Quality: 246.00 Length: 239
Ratio: 2.103 Gaps: 15
Percent Similarity: 48.954 Percent Identity: 37.238

alignment_block:
US-09-215-435-225 x AC037441/rev ..
Align seg 1/1 to reverse of: AC037441 from: 1 to: 162795

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
159302 ATGGGTGGACAATGAGCTGTGTACAGCAGCAGCAGCTGTACTGGGCTCAT 159253
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
159252 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 159203
34 laLeuLeuAspGluAspThrLeuPheCysGln.....GlyLeu 46
159202 CCCTCTGGACGAGGACACCCCTCTTTGCCAGTAGTCAGGCGAGCGCTTG 159153
47 GluValPheThrProGlyLeuGlyAsnIleGlyCysLysValValProAs 63
159152 GAGGCGGAGGCGACAGAGCT..... 159132
63 pCysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleValL 80
159131ACCAGGCAG..... 159123
80 yspheProGlyAlaValAspGlyAlaThr...TyrIleLeuValMetVal 95
159122GTAGATGGCAGGAGGGGTATATC.....CAG 159096
96 AspProAspAlaProSerArgAlaGluPro.....ArgGln... 107
159095 AACCAAGACGACCATCT...CCTGAGCCTGATTTATGGGAGGCAGAA 159049
108ArgPheTrpArgHis.....TrpLeuValThrAspIleLysG 120
1.....:|||||

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

[illegible]

```
FEATURES
source
misc_f
```

Thu Nov 15 10:52:08 2001

```

97 ProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis... 112
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49428 CCTGATTTTATGGAGGAGGACACACAGTGTCTAG...ATCTGGAGACACACC 49382
113 ....TrpLeuValThrAspIleLeuGlyAlaAspLeuLysLysGlyLysI 128
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49381 TTGATGGAGGATGGCA.....CTGAGAAAGGA...T 49353
128 leGlnGlyGlnGluLeuSerAlaTyGrGlnAlaProSerProAlaHis 144
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49352 TTAGAACAGACTCTCTCTGCTTCCAGATGCACCTCTCCAGGAGCA 49303
145 SerGlyPheHisArgTyGrGlnPheValTyLeuGlnGluGlyLysVa 161
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49302 GGAGGGCTGTGTGTGG..... 49285
161 lIleSerLeuLeuProLysGluAsnLysThrArgGlySerTrpLysMeta 178
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49284 ....TCGGCTCTGCTTGGGCCCTCCAAACCTGGGGGTCT..... 49249
178 spArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerThrGln 194
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49248 .....TTGTGGGGCGGAGGCTGAGACC... 49225
195 PheMetThrGlnAsnTyGrGlnAspSerProThrLeuGlnAla..... 208
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49224 ...CTCAGAGAGACACCCAGAGAGACCATCCCTCCCTGCAGCCTGAAT 49178
209 ....ProArgGluArgAlaSerGluProLysHisLysAsnGln 221
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
49177 GCGCCCTCTCCCGAGACCCCAACAAAGAGGCATTTGANTGAA 49135
seq_name: gb_ro2:RNPBP

```

```

seq_documentation_block: 1037 bp mRNA ROD 17-FEB-1997
LOCUS RNPBP
DEFINITION R.norvegicus mRNA for phosphatidylethanolamine-binding protein.
ACCESSION X71873
VERSION X71873.1 GI:510338
KEYWORDS epididymal protein; pbp gene; phosphatidylethanolamine-binding
protein.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1037)
AUTHORS Perry,A.C., Hall,L., Bell,A.B. and Jones,R.
TITLE Sequence analysis of a mammalian phospholipid-binding protein from
testis and epididymis and its distribution between spermatozoa and
extracellular secretions
JOURNAL Biochem. J. 301 (Pt 1), 235-242 (1994)
MEDLINE 94311839
REFERENCE 2 (bases 1 to 1037)
AUTHORS Hall,L.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-1993) L. Hall, University of Bristol, Dept of
Biochemistry, School of Medical Sciences, University Walk, Bristol
BS8 1TD, UK
FEATURES
source
1..1037
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/dev_stage="adult male"
/tissue_type="epididymis and liver"
/clone_lib="cDNA"
/clone="18 clones"
1..28
29..592
29..592
/gene="pbp"
5'UTR
gene
CDS
/gene="pbp"

```

```

clone_end:SP6
vector_side:left
misc_feature 436..2289
/notes="assembly_fragment"
misc_feature 2390..4997
/notes="assembly_fragment"
misc_feature 5098..64677
/notes="assembly_fragment"
misc_feature 64778..68151
/notes="assembly_fragment"
misc_feature 68252..72655
/notes="assembly_fragment"
misc_feature 72756..78893
/notes="assembly_fragment"
misc_feature 78994..84664
/notes="assembly_fragment"
misc_feature 84765..90853
/notes="assembly_fragment"
misc_feature 90954..98425
/notes="assembly_fragment"
misc_feature 98526..106371
/notes="assembly_fragment"
misc_feature 106472..115673
/notes="assembly_fragment"
misc_feature 115774..127138
/notes="assembly_fragment"
misc_feature 127239..138436
/notes="assembly_fragment"
misc_feature 138537..149766
/notes="assembly_fragment"
misc_feature 149867..167340
/notes="assembly_fragment"
misc_feature 167441..169645
/notes="assembly_fragment"
clone_end:r7
vector_side:right
BASE COUNT 44253 a 41074 c 39720 g 42960 t 1638 others
ORIGIN
alignment_scores:
Quality: 246.00 Length: 231
Ratio: 2.067 Gaps: 12
Percent Similarity: 51.515 Percent Identity: 36.364
alignment_block:
US-09-215-435-225 x AC084081/rev ..
Align seg 1/1 to reverse of: AC084081 from: 1 to: 169645
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuLeuLeu 17
|||||
49659 ATGGGTGGACAAATGAGGTGTGTACACAGCAGCAGTGTACTTGGGTCTCAT 49610
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
49609 GATGGTGGTCACTGGAGCAGGAGGATGAGAACAGCCCGTGTGCCATGAGG 49560
34 laLeuLeuAspGluAspThrLeuPheCysGln.....GlyLeu 46
|||||
49559 CCCTCTGGAGGAGGACACCCCTTTTGGCCAGCACTAGTCAGGCGAGGCTTG 49510
47 GluValPheTyGrProGluLeuGlyAsnIleGlyCysLysValValProAs 63
|||||
49509 GAGCGCGGA.....GGCACAGAGGCTACACAGCAGGTAGATGG 49472
63 pCysAsnAsnTyGrGlnLysIleThrSerTrpMetGluProIleValL 80
|||||
49471 CAGGGAGGGGTATATCCAGAAC..... 49450
80 ysPheProGlyAlaValAspGlyAlaThrTyIleLeuValMetValAsp 96
|||||
49449 .....CAAGGAGCACCA.....TCTCCTGAG 49429

```

```
/codon_start=1
/product="phosphatidylethanolamine-binding protein"
/protein_id="CAA50708.1"
/db_xref="GI:510339"
/db_xref="SWISS-PROT:P31044"
/translation="MAADISQWAGPLSLQEVDEPPQHALRVDYGGVTVDELGKVLPTF
QVNRPSISWMDGLDPGLKLYTLVTDAPSRKDPKFRWHHFLVNMKGNDISSGTV
LSEYSGSPKPDGLHRYVWLVEQQLNCDEPILSNKSGDNRGRKFVSEFRKKYHL
GAPVAGTCFOAEWDDSVPLKLDQLAG"
3'UTR
variation
766
/note="in 6 clones"
BASE COUNT      244 a      299 c      290 g      204 t
ORIGIN
alignment_scores:
  Quality: 237.50      Length: 185
  Ratio: 2.284        Gaps: 6
  Percent Similarity: 56.216      Percent Identity: 32.973
alignment_block:
US-09-215-435-225 x RNPBP ..
Align seg 1/1 to: RNPBP from: 1 to: 1037
52  GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 68
|||||
134  GAGCTGGGCAAGTG.....CTGAGCGCCACCCAGGTCATGAATAG 174
68  gGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAlaVal 85
| : : : |||||
175  ACCAAGCACCATTCATGG.....GATGCCTTG 203
85  alAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 101
||| ||| ||||| : : : |||||
204  ATCCTGGGAAGCTCTACACCTGGTCTCTCACAGACCCGATGCTCCAGC 253
102  ArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIle 118
||| : : : ||||| ||| ||||| : : : ||||| : : : |||||
254  AGAAGGACCCCAATTCAGGAGTGGACCACTTCCTGGTGGTCAACAT 303
118  eLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerA 135
: ||||| ||| : : : : : ||| : : : |||||
304  GAAGGCACGACATTAGACTGGCACTGTC.....CTCTCCG 341
135  laTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGln 151
||| : : : ||||| ||| : : : |||||
342  AATACGTGGCTCCGACCTCCCAAGACACAGGCTGCACCGCTACGTC 391
152  PhePheValTyrLeuGlnGluGlyLysValIleSer.....Le 164
: : : : : ||||| ||||| : : :
392  TGGCTGGTGTATGAGCAGGAGCAGCCTCTGAACCTGTGACGACCCATCT 441
164  uLeuProLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheL 181
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
442  CAGCAACAGTCTGGAGACACCGCGCAAGTTCAGGTGGAGTCTCTCC 491
181  euAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThr 197
: : : : : ||||| ||| ||| : : : |||
492  GCAAGAAGTACCACCTGGGAGCCCGCGTGGCGGACGTCCTCCAGGCA 541
198  GlnAsnTyrGlnAspSer.....
: : : : : |||||
542  GAG...TGGGATGACTCTGTGCCAAGCTGCACGATCAGTCAGTGGGAA 588
204  .....ProThrLeuGlnAlap 209
|||||
589  GTAGGGCGCTGCAGAGCCCGCAGCCCGGGGAGCCACACAGTACAGTCAA 638
209  roArg 210
|||
```

639 GTCGT 643

seq_name: gb_pat2:E05646

seq_documentation_block: 1047 bp RNA PAT 29-SEP-1997
LOCUS E05646 cDNA encoding rat hippocampal cholinergic neurostimulating
DEFINITION peptide, HCNP.
ACCESSION E05646
VERSION E05646.1 GI:2173833
KEYWORDS JP 1993268966-A/1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Rattus.

REFERENCE 1 (bases 1 to 1047)

Authors Tofuji, N., Tojo, S., Kojima, S., Ueki, Y., Nishihara, N., Fukushima, N.,
Irie, T., Ono, K., Agui, H. and Kojima, Y.
TITLE NEUTROPHIC FACTOR-RELATED POLYPEPTIDE AND GENE USEFUL THEREFOR
JOURNAL Patent: JP 1993268966-A 1 19-OCT-1993;
SUMITOMO PHARMACEUT CO LTD, KOJIMA YUKIO, YAMAMOTO MASAHIKO

COMMENT OS Rattus norvegicus (rat)

PN JP 1993268966-A/1

PD 19-OCT-1993

PF 25-APR-1992 JP 1992131703

PR 27-APR-1991 JP 91P 124688

PI TOFUJI NAOKI, TOJO SHINICHIRO, KOJIMA SHINICHI, UEKI YASUYUKI,

PI NISHIHARA NORIO, FUKUSHIMA NOBUYUKI, IRIE TSUNEMASA, PI ONO

KEIICHI,

PI AGUI HIDEO, KOJIMA YUKIO

PC C12N15/12,A61K37/02,A61K37/02,C07K7/06,C07K13/00,C12N1/21, PC

C12N5/10,

PC (C12N1/21,C12R1:19),(C12N5/10,C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: tissue-type-hippocampus of brain;

CC *source: clone-A010-12;

CC Feature is identified by similarity;

FH Key Location/Qualifiers

FH 5'UTR <1..25

FT CDS 26..589

FT /product='hippocampal cholinergic FT

FT neurostimulating peptide

FT 3'UTR 590..1047

FT polyA_signal 987..992

FT polyA_site 1008..1047

FT mutation replace(427..429,'AAG')

FT /note='hippocampal cholinergic FT

FT neurostimulating peptide

FT derived from clone A01-1'

FEATURES Location/Qualifiers

source 1..1047

/organism='Rattus norvegicus'

/db_xref='taxon:10116'

BASE COUNT 257 a 297 c 289 g 204 t

ORIGIN

alignment_scores:

Quality: 237.50 Length: 185

Ratio: 2.284 Gaps: 6

Percent Similarity: 56.216 Percent Identity: 32.973

alignment_block:

US-09-215-435-225 x E05646 ..

Align seg 1/1 to: E05646 from: 1 to: 1047

52 GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 68

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rwnematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is C37C3, 200 bp overlap, 3' lies in a gap. Actual
start of this cosmid is at base position 197 of CELF40A3; actual
end is at 41155 of CELF40A3

NOTES:

Coding sequences below are predicted from computer analysis, using
the program GeneFinder (P. Green and L. Hillier, ms in preparation).

FEATURES
source

Location/Qualifiers
1. 41155
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="v"
/clone="F40A3"
complement(156..1605)
/gene="F40A3.5"
complement(join(156..216,268..662,714..787,864..981,
1094..1281,1357..1431,1476..1605))

/note="F40A3.5"
/note="contains similarity to a protein kinase domain"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65324.1"
/db_xref="GI:2291201"

/translation="MELDHTPPSVLNDNCASVYTPVATVIAMSGLYLLAIFYECKK
SKKCMPSDSIYPYQKRLTQLERELFDVRAHSYFFAFLKNLYLDEESIEVDQFQIG
QTADGFTFGGVFPKTRNFNAKVTAVKISPIVSKSISLLEADRLSKLDHNPILIR
LIAVSQSLSEVFPMLALELPGTLADYQFKVREKDDSDERSPIQLKMLSLYQVS
QALKYIHSQIDFEGQELTHGRIFTRNLVTEPDLRKEVKLGDFGADPMGLEIYSPPII
AYMPPEILCAERIPPHRPENDVMFGVFIWECLTLGAQPHFRKSVEEIKKSFRLPDR
GLSCPPTCPLDV"
6740..7594
/gene="F40A3.4"
join(6740..6827,7331..7394,7450..7594)
/gene="F40A3.4"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65323.1"
/db_xref="GI:2291200"

/translation="MSRMSDYSSSIREEQAILNGLLGLFADGIYDNDVTRKHQ
TVDCGIYRMVALLIVIALIIIAVFLICEFHCRANNEERSPLHSNRKSHPIEI"
8506..9283
/gene="F40A3.3"
join(8506..8600,8667..8781,8828..9283)
/gene="F40A3.3"
/note="Similar to phosphatidylethanolamine binding
protein; coded for by C. elegans cDNA yk106a12.5; coded
for by C. elegans cDNA yk106a12.3"
/codon_start=1

gene
CDS

/protein_id="AAB65322.1"
/db_xref="GI:2291199"
/translation="MVLVTRSLLPALFFASRAFFAAATTSARFQRLATMAAEAFTK
HEVIPDLASNPFSKVSXVFNSEGVANLGNVLTPTQKDTPEVKWDAPGLVYLIYK
TDDAPSRKEHTYREWHHLVNVINPGNDIAKGDTLSEYICAGPPKTKLHRYLYIYK
QSGRIEDAHEGRLTNTSGDKRGGWKAADFVAKHKLGAPEFGNLFQAEYDDYVILNKQ
LGA"
9824..12167
/gene="F40A3.2"
join(9824..9917,10735..11058,11837..12034,12091..12167)
/gene="F40A3.2"
/note="contains similarity to EGF-like domains"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65320.1"
/db_xref="GI:2291197"

/translation="MLRLLLSAILVIVACVLRRKKKPCINGTPEGDRCYCIEGWTC
LCHRMVCSGFERLQNGSCIECELGWAGTGDIDCHNGMNPYDLTCTCTVPSYSGK
YCEIADTKDIYKWNNTSSIGPILITIIPLALLIYCCBEHFAKKRRLRVEOHLNGV
YITTRGQVDRKLVKGLLEASDDEKSVKVKIAADSGEKDKDEEVEVIMRNSATTS
EKKALLADPADM"
complement(23146..24752)
/gene="F40A3.6"
complement(join(23146..23311,23512..23597,23644..23841,
23885..24052,24594..24752))
/gene="F40A3.6"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65325.1"
/db_xref="GI:2291202"

/translation="MSEVLGWHNRKPAAGSRGNSFFDRDIVPGGAHTKSYESYVKNIF
HPTNRUKADQLVDRHGLSATSLRAEREGRSALDDKYDRHEFFEKMPLESYAE
LGLKNPAASNYRSFOYEPKSSSEYFAASYEAERERRKKDLERELASVDKEDYDFTG
EPKNLYSSISYRPLSSYKSRDEDEYEKKNRKLLIDYKDFDSSPLSPYLYQSHPYV
RSQDSRIIGSNFVQLTSRPKDKFISKIETLAWRDMRPY"
24968..26143
/gene="F40A3.1"
join(24968..25107,25199..25454,25498..25634,26041..26143)
/gene="F40A3.1"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65321.1"
/db_xref="GI:2291198"

/translation="MSEIVRLPEETOETEMPLHNAFARGARKQDAGWAKNEGKTFR
VLIKFEIFOINDWELIGSRYSWKNEVLRCRHKMDKVAIRWHLSCTCEHGVCDGDTN
VDYATRAONKTEMVVRNFTVVOIPGLOKTKFLRLGLTYDSRKTQLEAFSPKID
VDWRQVTRDGCNNFVVRNAMSINKMRKLERFCRLDKDHFVDCFTNL"
complement(26520..30751)
/gene="F40A3.7"
complement(join(26520..26710,27914..28026,28197..28358,
28415..28609,29618..29761,29814..29935,30397..30616,
30669..30751))
/gene="F40A3.7"
/note="contains similarity to seven-transmembrane
receptors"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65319.1"
/db_xref="GI:2291196"

gene
CDS

/translation="MSANFSLCLTDEOQRMSGTNLEYNLQRYEYPALATFGLGNVL
NLTVLLNRSRNSRANSLAVAFADITFLFPNLIANYSFETFWFMYLHSHK
HLISLANWSSVAHWCVIAGCDRLGIONPLVARTRWKKLPVLTIVTCGLIT
CYOHFEYCLVRSYRETQLYSRCLPNNAEKWCHRNPFQSYQSFIACKLAHLIFL
MIILPITALLLENLTLWAGRKRLSIGDNADNRQNDVHMOKTHERVTLTVTFI
VTMTLTNGPSALVHLVNYATHEELYDLTMTISSTLVTCGRASNFILFCLSGKHFRLRL
LKLTRKKINRKTIDSTIGTSLVHTTKLTSTASRRTSMPIOEKKRSSCAPLVAKKNALST
VETIQPLINGNRKGF"
13599 a 7323 c 6798 g 13435 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 230.00

Length: 159

Ratio: 2.447 Gaps: 7
Percent Similarity: 59.119 Percent Identity: 37.107
alignment_block:
US-09-215-435-225 x CELF40A3 ..
Align seg 1/1 to: CELF40A3 from: 1 to: 41155
49 PheTyrProGluLeuGlyAsnIleGlyCysValValProAspCysAs 65
8822 TTTCAGGCTAATCTTGGAAACGTT.....CTCACTCCACCC..... 8857
65 nAsnTyrArgGlnLysIleThrSerTrpMetGluProIleValLysPheP 82
8858CAAGTAAAGGATAC.....CCAGAAAGTAAAGTGG..... 8887
82 roGlyAlaValAspGlyAlaThrTrpIleLeuValMetValAspProAsp 98
8888 ..GATGCTAGCCAGGACGACCTCTACACTCTCATCAAGACGGATCCAGAT 8935
99 AlaProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuVa 115
8936 GCTCCATCTCGCAAGAGCCACCTACCGAGACTGGCATCACTGGCTCGT 8985
115 lThrAspIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnG 132
8986 CGTCAACATTCAGGAACGACATGCCCCAAA.....GGAGACA 9023
132 luLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHis 148
9024 CTCTTTCCGAGTACATCGCGCGGACCACTACCAACCACTGGACTCCAT 9073
149 ArgTyrGlnPhePheValTyrLeuGlnGlyLysValIleSer..... 163
9074 CGTTAGCTTTACTGATCTACAAAGCAATCTGGACCATCGAAGATCCGA 9123
164LeuLeuProLysGluAsnLysThrArgGlySerTrpLysM 177
9124 GCACGGACGCTCTAACCAACACTTCGGAGACAAAGAGGAGGATGGAAG 9173
177 eTAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerThr 193
9174 CTGCGGATTTCTGCGCAAGACAAAGCTCGGTCTCCAGTCTTCGGAAT 9223
194 GlnPheMetThrGlnAsnTyrGlnAsp 202
9224 CTGTTCCAGCGCGAG...TACGACGAC 9247
seq_name: gb_htg1:AC006915

seq_documentation_block:
LOCUS AC006915 134095 bp DNA HTG 24-FEB-1999
DEFINITION Caenorhabditis elegans clone Y97E10y, *** SEQUENCING IN PROGRESS
***, 1 unordered pieces.
ACCESSION AC006915
VERSION AC006915.1 GI:4263433
KEYWORDS HTG: HTGS, PHASE1
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 134095)
Waterston,R.H.
The sequence of Caenorhabditis elegans clone
Unpublished
JOURNAL 2 (bases 1 to 134095)
REFERENCE Waterston,R.H.
AUTHORS Direct Submission
TITLE Submitted (24-FEB-1999) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT * NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 134095: contig of 134095 bp in length.
FEATURES
source
1. 134095
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="Y97E10y"
BASE COUNT 43230 a 24499 c 23702 g 42664 t
ORIGIN
alignment_scores:
Quality: 230.00 Length: 159
Ratio: 2.447 Gaps: 7
Percent Similarity: 59.119 Percent Identity: 37.107
alignment_block:
US-09-215-435-225 x AC006915 ..
Align seg 1/1 to: AC006915 from: 1 to: 134095
49 PheTyrProGluLeuGlyAsnIleGlyCysValValProAspCysAs 65
99439 TTTCAGGCTAATCTTGGAAACGTT.....CTCACTCCACCC..... 99474
65 nAsnTyrArgGlnLysIleThrSerTrpMetGluProIleValLysPheP 82
99475CAAGTAAAGGATAC.....CCAGAAAGTAAAGTGG..... 99504
82 roGlyAlaValAspGlyAlaThrTrpIleLeuValMetValAspProAsp 98
99505 ..GATGCTAGCCAGGACGACCTCTACACTCTCATCAAGACGGATCCAGAT 99552
99 AlaProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuVa 115
99553 GCTCCATCTCGCAAGAGCCACCTACCGAGAGTGGCATCACTGGCTCGT 99602
115 lThrAspIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnG 132
99603 CGTCAACATTCAGGAACGACATTCGCCAAA.....GGAGACA 99640
132 luLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHis 148
99641 CTCTTTCCGAGTACATCGCGCGGACCACTACCAACCACTGGACTCCAT 99690
149 ArgTyrGlnPhePheValTyrLeuGlnGlyLysValIleSer..... 163
99691 CGTTAGCTTTACTGATCTACAAAGCAATCTGGACCATCGAAGATCCGA 99740
164LeuLeuProLysGluAsnLysThrArgGlySerTrpLysM 177
99741 GCACGGACGCTCTAACCAACACTTCGGAGACAAAGAGGAGGATGGAAG 99790
177 eTAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerThr 193
99791 CTGCGGATTTCTGCGCAAGACCAAGCTCGGTCTCCAGTCTTCGGAAT 99840
194 GlnPheMetThrGlnAsnTyrGlnAsp 202
99841 CTGTTCCAGCGCGAG...TACGACGAC 99864
seq_name: gb_htg4:AC012978
seq_documentation_block:
LOCUS AC012978 63551 bp DNA HTG 03-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
pieces.
ACCESSION AC012978

Thu Nov 15 10:52:08 2001

```

VERSION      AC012978.1  GI:6223345
KEYWORDS     HTG; HTGS_PHASE2.
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
              Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
              Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 63551)
AUTHORS      Adams,M. and Venter,J.C.
TITLE        Direct Submission
JOURNAL      Submitted (03-Nov-1999) Celera Genomics, 45 West Gude Drive,
              Rockville, MD, USA
COMMENT      This sequence was identified as CDM:10212016 by the submitter.
              For further information on this sequence you may e-mail to
              fly@celera.com.
              * NOTE: This is a 'working draft' sequence.
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.
              * the accession number will be preserved.

FEATURES             Location/Qualifiers
     source           1..63551
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT          18144 a 13546 c 14099 g 17762 t
ORIGIN

alignment_scores:
    Quality: 229.00      Length: 131
    Ratio: 2.516        Gaps: 5
    Percent Similarity: 69.466      Percent Identity: 40.458

alignment_block:
US-09-215-435-225 x AC012978/rev ..
Align seg 1/1 to reverse of: AC012978 from: 1 to: 63551

76 GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 92
23000 GAGCGGTGGTGAATGGGAG...GCGACGCCAATAAGTGTACACCT 22954
92 uValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgp 109
22953 CTGCATGCCATCGGATCGCCAGTCGCAAGGATCCCAAGTTTAGGG 22904
109 heTrpArgHisTrpLeuThrAspIleLysGlyAlaAspLeuLysLys 125
22903 ACTGGCACCATGGTGGTGGGCAACATACCCGCTGGAGATGTCGCCAAG 22854
126 GlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProPr 142
22853 GCGGAGGT.....CTCTCCGCTACGTGGGATCGCGGCTCC 22816
142 oAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGlu 159
22815 ACCAGACACCGGACTCCATCGTTAGCTTTCTGATCTACGACGCGGT 22766
159 lyLysValIle.....SerLeuLeuProLysGluAsn...LysThr 171
22765 GCAGGCTCACATTGCGAGAGAGCGACTGCCCAATAACAGCGGATGGA 22716
172 ArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyCl 188
22715 CGCGGTGGCTTCAAAATCGCGAGTTCGCCAAGAAGTAGCCCTCGGCA 22666
188 uproGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAsp 202
22665 TCCCATTCGCGGTAACTCTTACAGGCTGAG...TAGCACGAC 22626

seq_name: gb_inl:AC005891
seq_documentation_block:
LOCUS       AC005891 136933 bp DNA INV 30-OCT-1998

```

```

DEFINITION     Drosophila melanogaster, chromosome 2L, region 33F3-34A2, pl clones
DS08479 and DS07153, complete sequence.
AC005891 AC005108 AC005109 AC005110
AC005891.1 GI:3810575
HTG.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 136933)
AUTHORS        Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
              Butenhoff,C., Chavez,C., Chew,M., Ciesiolka,L.,
              Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
              Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
              Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
              Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
              Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
              Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
              Zieran,L.L. and Rubin,G.M.
TITLE          Sequencing of Drosophila chromosome 2L, region 33F3-34A2
JOURNAL        Unpublished (1998)
AUTHORS        2 (bases 1 to 136933)
              Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
              Butenhoff,C., Chavez,C., Chew,M., Ciesiolka,L.,
              Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
              Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
              Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
              Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
              Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
              Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
              Zieran,L.L. and Rubin,G.M.
TITLE          Direct Submission
JOURNAL        Submitted (30-OCT-1998) Berkeley Drosophila Genome Project, MS
              64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
              Berkeley, CA 94720, US
COMMENT        Sequence submitted by:
              Lawrence Berkeley National Laboratory, MS 64-121
              Berkeley, CA 94720
              For further information about this sequence, including its location
              and relationship to other sequences, please visit our sequence
              archive Web site (http://www.fruitfly.org/sequence/) or send email
              to bdg@fruitfly.berkeley.edu.
              pl library locations: 89-31, 75-49.
FEATURES       Location/Qualifiers
               1..136933
               /organism="Drosophila melanogaster"
               /strain="v2: cn bw sp"
               /db_xref="taxon:7227"
               /chromosome="2L"
               /map="33F3-34A2"
               /clone_lib="pl library, partial Sau3A in pAd10sacBII"
               /clone="Pls DS08479 (D303) and DS07153 (D305)"
               /note="DS08479 (D303) and DS07153 (D305) were completed as
               a project. DS08479 extends from pl end at bp 1 to pl end
               at bp 83797. DS07153 extends from pl end at bp 51,253 to
               pl end at bp 136,933."
BASE COUNT     37380 a 28396 c 29033 g 42124 t
ORIGIN

alignment_scores:
    Quality: 229.00      Length: 131
    Ratio: 2.516        Gaps: 5
    Percent Similarity: 69.466      Percent Identity: 40.458

alignment_block:
US-09-215-435-225 x AC005891 ..
Align seg 1/1 to: AC005891 from: 1 to: 136933

76 GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 92

```

```
||||| |||||::: ||| ::: ||| ||
24363 GAGCCGTGCGTGAATGGGAG...GCGGACGCCAATAAGCTGTACACCGT 24409
92 uValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgP 109
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24410 CTGCATGACCGATCCGATGCGCGATGCGCCCGAGTCGCAAGGATCCCAAGTTAGGG 24459
109 heTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 125
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24460 AGTGGCACCATTTGGCTGGTGGCAACATACCGGTGGAGATGTCGCCAAG 24509
126 GlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProPr 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24510 GCGGAGGT.....CTCTCCGCCTACGTGGGATCCGGGCTCC 24547
142 oAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGluG 159
| ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24548 ACCAGACCGGACTCCATCGTTACGTTTCTGATCTACGAGCAGCGGT 24597
159 lYlYSValIle.....SerLeuLeuProLysGluAsn...LysThr 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24598 GCAAGCTCACATTCGACGAGAGCGACTGCCCAATAACAGCGGAGATGGA 24647
172 ArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyG 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24648 CGCGTGGCTTCAAAATCGCCGAGTTCGCCAAGAAGTACGCCCTCGGCAA 24697
188 uProGluAlaSerThrGlnPheMetThrClnAsnTyrGlnAsp 202
:|||| ||||:|||||:|||||:|||||:|||||:|||||:|||||:
24698 TCCCATTCGCCGTAACCTCTACGAGGCTGAG...TACGACGAC 24737
```

THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-225 to: N_Geneseq_0601.* out_format : pfs
 Date: Nov 15, 2001 4:29 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=framet.p2n.model -DRV=xlp
 -Q/cgml_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
 -DB=N_Geneseq_0601 -QPM=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
 -TRANS=human40.cdi -LISP=45 -DOCLIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09215435_@CGML_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225
 Query length: 227
 Database: N_Geneseq_0601.*
 Database sequences: 730101
 Database length: 313950809
 Search time (sec): 1043.840000

score_list:	Sequence	Strd Orig	zscore	Escore Len	! Documentation
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480		1224.00	2543.16	1.1e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA97660		1224.00	2541.04	1.4e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA90013		1224.00	2541.04	1.4e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA90013		1224.00	2541.04	1.4e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA94490		1224.00	2540.14	1.6e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA94482		1220.00	2531.34	5.0e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA978426		1218.00	2526.07	9.9e-133
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA97539		1215.00	2521.14	1.9e-132
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA975582		1213.00	2517.02	3.1e-132
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA978431		1198.00	2485.01	1.9e-130
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA96350		1189.00	2466.89	2.0e-129
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA96350		652.00	1349.31	3.5e-67
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA940582		237.50	471.50	2.7e-18
	/cgml_9/gcgdata/geneseq/geneseq/NA1992.DAT:AA930001		237.50	471.50	2.7e-18
	/cgml_9/gcgdata/geneseq/geneseq/NA1994.DAT:AA938685		208.50	407.19	1.0e-14
	/cgml_9/gcgdata/geneseq/geneseq/NA1992.DAT:AA930002		208.50	407.19	1.0e-14
	/cgml_9/gcgdata/geneseq/geneseq/NA1994.DAT:AA938686		207.50	409.14	8.1e-15
	/cgml_9/gcgdata/geneseq/geneseq/NA1995.DAT:AA980734		201.50	398.73	3.1e-14
	/cgml_9/gcgdata/geneseq/geneseq/NA1991.DAT:AA941482		182.00	362.72	3.1e-12
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA960884		182.00	362.72	3.1e-12
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA942966		164.00	325.01	3.9e-10
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA960883		164.00	321.80	6.5e-10
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA948359		164.00	320.80	6.7e-10
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA948359		164.00	319.71	7.7e-10
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA928490		164.00	319.71	7.7e-10
	/cgml_9/gcgdata/geneseq/geneseq/NA1997.DAT:AA960139		148.00	285.29	6.4e-08
	/cgml_9/gcgdata/geneseq/geneseq/NA1997.DAT:AA960139		147.50	290.63	5.1e-08
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA960868		145.50	286.37	3.5e-08
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA960868		140.50	273.19	3.0e-07
	/cgml_9/gcgdata/geneseq/geneseq/NA1997.DAT:AA960140		121.00	238.84	2.5e-05
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA93828		90.50	162.01	0.713
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA987678		90.50	173.83	0.1034
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA939625		89.50	165.06	0.3183
	/cgml_9/gcgdata/geneseq/geneseq/NA1998.DAT:AA939625		89.50	165.06	0.3183
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA981032		89.50	158.41	0.7475
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA985554		88.00	157.15	0.8787
	/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA987700		88.00	82.59	1.2e+04
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA9201425		87.50	142.45	5.79
	/cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA927717		87.50	140.53	7.41
	/cgml_9/gcgdata/geneseq/geneseq/NA1993.DAT:AA951531		87.50	140.53	7.41
	/cgml_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA916969		87.50	140.53	7.41

/cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA953434 + 86.00 152.20 1.66
 /cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA977512 + 85.00 136.86 11.86
 /cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA902705 + 83.50 153.16 1.47
 /cgml_9/gcgdata/geneseq/geneseq/NA1996.DAT:AA934817 + 83.50 152.55 1.58
 /cgml_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA978058 + 83.50 143.90 4.81

seq_name: /cgml_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480

seq_documentation_block:
 ID AAF94480 standard: cDNA; 681 BP.

XX AAF94480;

DT 04-JUN-2001 (first entry)

DE Human hydrophobic domain containing protein clone HP03880 cDNA #104.

XX Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
 KW antianemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
 KW cytosolic; gene therapy; autoimmune disorder; multiple sclerosis;
 KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
 KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
 KW behavioural characteristic; immune response; ss.

XX Homo sapiens.

XX WO200112660-A2.

XX 22-FEB-2001.

XX 10-AUG-2000; 2000WO-JP05356.

XX 17-AUG-1999; 99JP-0230344.

XX 07-SEP-1999; 99JP-0252551.

XX 01-OCT-1999; 99JP-0281132.

XX 22-OCT-1999; 99JP-0301624.

XX 04-NOV-1999; 99JP-0313877.

XX (SAGA) SAGAMI CHEM RES CENT.

XX (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;

XX WPT; 2001-160059/16.

XX P-PSDB; AAB88590.

XX Human proteins with hydrophobic domains and the DNAs which encode them are useful for treating autoimmune disorders, burns and tumors and for screening novel pharmaceuticals -

XX Claim 3; Page 426-427; 518pp; English.

XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant, anti-HIV, neuroprotective, antianemic, vulnery, antiulcer, osteopathic, anti-inflammatory and cytostatic activities, and can be used in gene therapy. (I) can be used as pharmaceuticals and as antigens to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as probes for genetic diagnosis and gene sources for gene therapy or for producing (I) in large quantities. Cells containing (II) are used for the detection of ligands or receptors corresponding to membrane or secretory proteins and to screen small molecule novel pharmaceuticals. Antibodies directed to (I) can be used for the detection, quantification and purification of (I). Activities of (I) may include cytokine and cell proliferation/differentiation function, immune stimulating or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activity/inhibitory activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity and anti-inflammatory activity. (I) and (II) can be used to treat autoimmune disorders e.g. multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis, inflammatory bowel disease and tumours. (I) and (II) can also be used for wound healing, as nutritional sources or supplements e.g. as amino acid, carbon or nitrogen source, to effect metabolism, catabolism, anabolism,

CC processing and utilisation of dietary fat, protein, carbohydrate,
 CC vitamins and minerals, to effect behavioural characteristics, to affect
 CC appetite, and can act as antigens in vaccines to raise an immune response
 CC to the protein or another material cross-reactive with the protein.

XX Sequence 681 BP; 179 A; 186 C; 183 G; 133 T; 0 other;
 SQ

alignment_scores:
 Quality: 1224.00 Length: 227
 Ratio: 5.392 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x AAF94480 ..

Align seg 1/1 to: AAF94480 from: 1 to: 681

1 MetClyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 1 ATGGTTGGACAATGAGCTGGTCACAGCAGCACTGTACTGGTCTCAT 50
 17 tmetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
 51 GATGGTGGTCACTGGAGACGAGGATGAGACAGCCGCTGTGCCCATGAGG 100
 34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
 101 CCCCTCTGGAGGAGACCCCTCTTTTGGCCAGGGCCTTGAAGTTTCTAC 150
 51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
 151 CCAGATTGGGGAACATTTGGCTGCAAGGTTGCTTCCTGATTGTAACAATA 200
 67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
 201 CAGACAGAAGATCACTCTCGATGGAGCCGATAGTCAAGTCCCGGGGG 250
 84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
 251 CGTGGACGGCGCACTATATCTGTTGGTGGATGATCCAGATGCCCT 300
 101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
 301 AGCAGACAGAACCCAGACAGATCTGGAGACATTTGGTGGTAAACAGA 350
 117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnLeuS 134
 351 TATCAAGGGCGCGACCTGAAGAAAGGAAGATTTCAGGGCCAGAGTTAT 400
 134 exAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
 401 CAGCCTTACAGGCTCCCTCCCGCCAGGACACAGTGGCTTCCATCGCTAC 450
 151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProLy 167
 451 CAGTTCTTTGTCTATCTTCAGGAGAAAGTCATCTCTCTCTCTCCCA 500
 167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
 501 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTTCTGAACCGTT 550
 184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
 551 TCCACCTGGGGGAACCTTGAGCAGGACCCAGTTCATGACCCAGAACTAC 600
 201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
 601 CAGGACTCACCAACCTTCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAA 650
 217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
 651 GCACAAAACACGCGGAGGATAGTACTGCTGTC 681

seq_name: /cgnl_9/gcdata/geneseq/geneseq/NAL1999.DAT:AAx97660

seq_documentation_block:

ID AAX97660 standard; DNA; 826 BP.

XX AAX97660;

XX 13-SEP-1999 (first entry)

XX Extended human secreted protein coding sequence, SEQ ID NO. 124.

DE Secreted protein; human; cytokine; cellular proliferation; cell movement;
 XX cellular differentiation; immune system regulator; anti-inflammatory;
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
 KW genetic disease; ss.

XX Homo sapiens.

XX WO9931236-A2.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-IB02122.

XX 10-AUG-1998; 98US-0096116.

XX 17-DEC-1997; 97US-0069957.

XX 09-FEB-1998; 98US-0074121.

XX 13-APR-1998; 98US-0081563.

XX (GEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI; 1999-385906/32.

XX P-PSDB; AAY35976.

XX New isolated human secreted proteins

XX Claim 1; Page 255; 516pp; English.

XX This sequence represents an extended human secreted protein coding
 CC sequence of the invention. The secreted proteins can be used in treating
 CC or controlling a variety of human conditions. The secreted proteins may
 CC act as cytokines or may affect cellular proliferation or differentiation
 CC or may act as immune system regulators, haematopoiesis regulators, tissue
 CC growth regulators, regulators of reproductive hormones or cell movement
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
 CC tumour inhibition activity. The DNAs can be used in forensic procedures
 CC to identify individuals or in diagnostic procedures to identify
 CC individuals having genetic diseases resulting from abnormal expression of
 CC the genes corresponding to the extended cDNAs. They are also useful for
 CC constructing a high resolution map of the human chromosomes. They can
 CC also be used for gene therapy to control or treat genetic diseases.

XX Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;
 SQ

alignment_scores:

Quality: 1224.00 Length: 227

Ratio: 5.392 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x AAX97660 ..

Align seg 1/1 to: AAX97660 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 15 ATGGTTGGACAATGAGGCTGGTTCACAGCAGCACTGTACTGGGTCTCAT 64

17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGTTGGTCACTGGAGACGAGGATGAGAACAACCCGCTGTGCCCATGAGG 114

34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuValPheTyr 50
|||||
115 CCCTCTGGACGAGGACACCCCTCTTTGCCAGGGCCTTGAAGTTTTCATAC 164

51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
165 CCAGAGTTGGGGAACATTGGCTGCAAGGTGTCTCTGATTGTAACAACATA 214

67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
215 CAGACAGAGATCACCTCTGATGGAGCGCATAGTCAAGTTCCCGGGG 264

84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCGTGGACGGCGCAACCTATATCCTGGTGATGGTGATCCAGATGCCCT 314

101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
315 ACCAGACAGACACCCAGACAGATTCCTGGACATTTGGCTGGTAACAGA 364

117 PileLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyLeuLys 134
|||||
365 TATCAAGGGCGCCGACCTCAAGAAAGGGAAGATTCAAGGCCAGGAGTTAT 414

134 eAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
|||||
415 CAGCTTACAGGCTCCCTCCCGCCAGCACAGTGGCTTCCATCGCTAC 464

151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProly 167
|||||
465 CAGTCTTGTCTATCTTCAGAAAGGAAGGTCACTCTCTCTTCCCAA 514

167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
515 GGAAACAAACCTCGAGGCTCTTGAAATGGACAGATTCTGAACCGTT 564

184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCCACCTGGGCAAGCTGAAGCAAGCACCAGTTCATGACCCAGAACATAC 614

201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
|||||
615 CAGGACTCACCAACCTCCAGCTCCAGAGAAAGGGCCAGCGAGCCCA 664

217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
665 GCACAAAACACCGGGGAGATAGCTGCCTGC 695

seq_name= /cgnl_9/gcdata/geneseq/geneseqn/NA2000.DAT.AAC00013

seq_documentation_block:

ID AAC00013 standard; cDNA; 826 BP.

XX AC AAC00013;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein cDNA sequence #4.

XX KW Human; secreted protein; 5' EST; expressed sequence tag; cDNA isolation;

XX KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX FT key

XX FT CDS

XX FT 15..698

XX FT /*tag= a

XX FT /product= "secreted protein"

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
DR P-PSDB; AAG00016.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Example 19; SEQ ID 7; 71pp + CD-ROM; English.
XX
CC The present sequence is a full length cDNA encoding a human
XX
CC secreted protein. The cDNA was obtained from a 5' EST using first and
XX
CC second strand synthesis procedures. 5' ESTs were prepared from total
XX
CC human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX
CC sequences usually correspond mainly to the 3' untranslated region
XX
CC (UTR) of the mRNA because they are often obtained from oligo-dT primed
XX
CC cDNA libraries. Such ESTs are not well suited for isolating cDNA
XX
CC sequences derived from the 5' ends of mRNAs and even in those cases where
XX
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
XX
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
XX
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
XX
CC are also used in diagnostic, forensic, gene therapy and chromosome
XX
CC mapping procedures. They are used to obtain upstream regulatory sequences
XX
CC and to design expression and secretion vectors.
XX
SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:

Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x AAC00013

Align seg 1/1 to: AAC00013 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGTCTCAT 64

17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGTTGGTCACTGGAGACGAGGATGAGAACAACCCGCTGTGCCCATGAGG 114

34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCCTCTGGACGAGGACACCCCTCTTTGCCAGGGCCTTGAAGTTTTCATAC 164

51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
165 CCAGAGTTGGGGAACATTGGCTGCAAGGTGTCTCTGATTGTAACAACATA 214

67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
215 CAGACAGAGATCACCTCTGATGGAGCGCATAGTCAAGTTCCCGGGG 264

84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCGTGGACGGCGCAACCTATATCCTGGTGATGGTGATCCAGATGCCCT 314

Thu Nov 15 10:52:09 2001

101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
 315 AGCAGAGCAGAACCCAGACAGAGATTCGGAGACATTCGCTGGTACACAGA 364
 117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
 365 TATCAAGGGCGCGACCTCAAGAAAGGGAAGATTTCAGGGCCAGGAGTTAT 414
 134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
 415 CAGCCTACCAAGGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 464
 151 GlnPhePheValTyrLeuGlnGluLysValIleSerLeuLeuProly 167
 465 CAGTTCCTTGTCTATCTTCAGGAAGGAAGGTCATCTCTCTCTCCCA 514
 167 sGluAsnLysThrArgLysSerTrpLysMetAspArgPheLeuAsnArgP 184
 515 GGAACAAACCTCGAGGCTCTTGGAAATGGACAGATTCTGAACCGTT 564
 184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
 565 TCCACCTGGGCGAACCTGAAGCAGACACCCAGTTCATGACCCAGAACTAC 614
 201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
 615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCGCCAGCGAGCCCAA 664
 217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
 665 GCACAAACACCGCGAGATAGTGCCTGC 695

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ42252

seq_documentation_block:
 ID AAZ42252 standard; cDNA; 826 BP.

XX AC AAZ42252;
 XX DT 01-FEB-2000 (first entry)
 XX DE Human phosphatidylethanolamine-binding protein encoding cDNA.
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification; ss.
 XX OS Homo sapiens.
 XX PN WO9953051-A2.
 XX PD 21-OCT-1999.
 XX PF 09-APR-1999; 99WO-IB00712.
 XX PR 09-APR-1998; 98US-0057719.
 XX PR 28-APR-1998; 98US-0069047.
 XX (GEST) GENSET.
 XX PA Dumas Milne Edwards J, Duclert A, Giordano J;
 XX PI WPI: 2000-038446/03.
 XX DR P-PSDB; AAY64647.
 XX Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 PS Example 21; Page 168-169; 837pp; English.
 XX AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to

CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
 CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 XX
 XX SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
 Quality: 1224.00 Length: 227
 Ratio: 5.392 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-215-435-225 x AAZ42252 ...
 Align seg 1/1 to: AAZ42252 from: 1 to: 826

1 MetGlyTyrThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
 15 ATGGGTGGACAATGAGCTGTGTACAGCAGCAGCTGTACTGGTCTCAT 64
 17 tMetValValThrGlyAspGluAspSerProCysAlaHisGluA 34
 65 GATGGTGTCTACTGGAGAGGAGATGAGACAGCCCGTGTCCCATGAGG 114
 34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
 115 CCTCTGTGACGAGGACACCTCTTTTGCAGGGCTTCAAGTTTCTTAC 164
 51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
 165 CCAGAGTTGGGAACATTTGGCTGCAAGGTTGTCTCTGATTTGAACACTA 214
 67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
 215 CAGACAGAAGATCACCTCTGGATGGAGCCGATAGTCAAGTTCGCGGGG 264
 84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
 265 CCGTGGACGGCGCAACCTATATCTGTGTGGATCCAGATGCCCT 314
 101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
 315 AGCAGAGCAGAACCCAGACAGATTCGAGACATTTGGCTGGTAACACA 364
 117 pIleLysGlyAlaAspLeuLysLysIleGlnGlyGlnGluLeuS 134
 365 TATCAAGGGCGCGACCTCAAGAAAGGGAAGATTTCAGGGCCAGGAGTTAT 414
 134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
 415 CAGCCTACCAAGGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 464
 151 GlnPhePheValTyrLeuGlnGluLysValIleSerLeuLeuProly 167
 465 CAGTTCCTTGTCTATCTTCAGGAAGGAAGGTCATCTCTCTCTCCCA 514
 167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184

|||||
515 GGAAACAAACTCGAGGCTCTGGAAATGGACAGATTCTGAACCGTT 564
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCCACCTGGCGCAACTGAAGCAAGCACCCAGTTCATGACCAGAACTAC 614
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGGCCAGCGAGCCCAA 664
217 shiLysAsnGlnAlaGluIleAlaLacys 227
|||||
665 GCACAAAAACCGGCGAGATAGTCCCTGC 695

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF94490

seq_documentation_block:

ID AAF94490 standard; cDNA; 897 BP.

XX

AC AAF94490;

XX

DT 04-JUN-2001 (first entry)

XX

DE Human hydrophobic domain containing protein clone HP03880 cDNA #114.

XX

KW Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;

KW antianaemic; vulnery; antiulcer; osteopathic; anti-inflammatory;

KW cytosatic; gene therapy; autoimmune disorder; multiple sclerosis;

KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;

KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;

KW behavioural characteristic; immune response; ss.

XX

OS Homo sapiens.

XX

PN WO200112660-A2.

XX

PD 22-FEB-2001.

XX

PF 10-AUG-2000; 2000WO-JP05356.

XX

PR 17-AUG-1999; 99JP-0230344.

PR 07-SEP-1999; 99JP-0252551.

PR 01-OCT-1999; 99JP-0281132.

PR 22-OCT-1999; 99JP-0301624.

PR 04-NOV-1999; 99JP-0313877.

XX

PA (SAGA) SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

XX

PI Kato S, Kimura T;

XX

XX WPI: 2001-160059/16.

DR P-PSDB; AAB88590.

XX

XX Human proteins with hydrophobic domains and the DNAs which encode them

PT are useful for treating autoimmune disorders, burns and tumors and for

PT screening novel pharmaceuticals -

XX

PS Claim 4; Page 442-444; 518pp; English.

XX

XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to

CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,

CC anti-HIV, neuroprotective, antianaemic, vulnery, antiulcer,

CC osteopathic, anti-inflammatory and cytostatic activities, and can be

CC used in gene therapy. (I) can be used as pharmaceuticals and as antigens

CC to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as

CC probes for genetic diagnosis and gene sources for gene therapy or for

CC producing (I) in large quantities. Cells containing (II) are used for

CC the detection of ligands or receptors corresponding to membrane or

CC secretory proteins and to screen small molecule novel pharmaceuticals.

CC Antibodies directed to (I) can be used for the detection, quantification

CC and purification of (I). Activities of (I) may include cytokine and cell

CC proliferation/differentiation function, immune stimulating or suppressing
CC activity, haematopoiesis regulating activity, tissue growth activity,
CC actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
CC wound healing, as nutritional sources or supplements e.g. as amino acid,
CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
CC processing and utilisation of dietary fat, protein, carbohydrate,
CC vitamins and minerals, to effect behavioural characteristics, to affect
CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein.
XX
SQ Sequence 897 BP; 226 A; 253 C; 234 G; 184 T; 0 other;

alignment_scores:

Quality: 1224.00 Length: 227

Ratio: 5.392 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x AAF94490

Align seg 1/1 to: AAF94490 from: 1 to: 897

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
99 ATGGGTGGACATAGGTGGTGCACGACGACACTGTACTGGGTCTCAT 148
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
149 GATGGTGGTCACTGGAGACGAGATGAGAACGCCGTGTGCCCATGAGG 198
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
199 CCCTCTTGGACGAGGACACCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 248
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAnty 67
|||||
249 CCAGAGTTGGGGAACATTGGCTCAAGGTTGTCTCTGATTGTAACAACTA 298
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
299 CAGACAGAAGATCACCTCCTGGATGCGGCGGATGATCAAGTTCCCGGGG 348
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
349 CCGTGGACGGCGCACTATATCTGTGTGTGGTGGATCCAGATGCCCT 398
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
399 AGCAGAGCAGACACCCAGACAGAGATTCTGGAGACATTGGCTGGTAACAGA 448
117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
449 TATCAGGGGGCGGACCTGAAGAAAGGGAAGATTTCAGGGCCAGAGTTAT 498
134 eAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
|||||
499 CAGCCTTACCAGGGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 548
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
549 CAGTTCTTTGTCTATCTTTCAGGAAGGAAAGTATCTCTCTCTCTCCCAA 598
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
599 GGAAACAAACTCAGAGGCTCTTGGAAATGGACAGATTCTGAACCGTT 648
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||

Thu Nov 15 10:52:09 2001

us-09-215-435-225.rng

```
649 TCACCTGGCGGAACCTGAAGCAAGCACCCAGTTCATGACCCAGAACTAC 698
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
|||||
699 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGGCCAGCGGCCCAA 748
217 sHisLysAsnGlnAlaGluLeuAlaLacys 227
|||||
749 GCACAAACACCGCGGAGATAGTCTGCCTGC 779

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC84882
seq_documentation_block:
ID AAC84882 standard; cDNA; 932 BP.
XX
AC AAC84882;
XX
DT 20-APR-2001 (first entry)
XX
DE Human SEC1 nucleic acid sequence (clone ID 3445452).
XX
KW SECX; cytostatic; gynecological; gene therapy; screening assay; human;
KW chromosomal mapping; forensic biology; cell proliferation; cancer;
KW cell differentiation; immune associated disorder; gestational disease;
KW SEC1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 113..781
FT FT /*tag= a
FT FT /product= "SEC1"
FT FT sig_peptide 113..181
FT FT /*tag= b
FT FT mat_peptide 182..778
FT FT /*tag= c
XX
XX W0200078802-A2.
XX
XX 28-DEC-2000.
XX
XX 23-JUN-2000; 2000WO-US17328.
XX
XX 23-JUN-1999; 99US-0140584.
XX
XX 20-JUL-1999; 99US-0144722.
XX
XX 16-SEP-1999; 99US-0154520.
XX
XX 22-JUN-2000; 2000US-0604286.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
PI Herrmann JL;
XX
XX WPI; 2001-071385/08.
XX
XX P-PSDB; AAB48368.
XX
XX Polynucleotides encoding SECX proteins useful for treating disease
PT characterized by an aberrant level of cell proliferation and/or
PT differentiation like cancer or immune associated disorders .
XX
XX Claim 3; Fig 1; 132pp; English.
XX
XX The invention relates to human SECX polypeptides and polynucleotides
XX encoding them. The SECX polypeptides can be expressed by standard
XX recombinant methodology. The SECX polypeptides are useful for treating
XX or preventing a SECX-associated disorder. The invention is useful in
XX screening assays; detection assays (e.g. chromosomal mapping, cell and
XX tissue typing, forensic biology); predictive medicine (diagnostic assays,
XX prognostic assays, monitoring clinical trials, and pharmacogenomics); and
XX methods of treatment (e.g. therapeutic and prophylactic), especially
XX disorders characterized by aberrant cell proliferation and/or
XX differentiation like cancer or immune associated disorders or gestational
XX disease. The present sequence represents a SEC1 nucleic acid sequence.
CC
```

```
XX
SQ Sequence 932 BP; 244 A; 257 C; 244 G; 187 T; 0 other;

alignment_scores:
Quality: 1220.00 Length: 227
Ratio: 5.374 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.559

alignment_block:
US-09-215-435-225 x AAC84882 ..
Align seg 1/1 to: AAC84882 from: 1 to: 932

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
113 ATGGGTTGGCAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 162
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
163 GATGGTGTCTACTGGAGCAGGAGATGAGAACAGCCCGTGTGCCCATGAGG 212
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
213 CCCTCTTGGAGCAGGACACCCCTCTTTGCCAGGGCCTTGAAGTTTCTAC 262
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
263 CCAGAGTTGGGGAAACATTGGCTGCAAGGTTCTCTGATTGTAACAACTA 312
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
313 CAGACAGAAGATCACCTCTGATGGAGCCGATAGTCAAGTTCCCGGGGG 362
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
363 CCGTGGAGCGCGCAACCTATATCTCTGGTGGTGGATCCAGATGCCCT 412
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
413 AGCAGAGCAGACAGCCAGACAGAGATTCTGGAGACATTGGCTGGTAACAGA 462
117 pileLysGlyAlaAspLeuLysLysLysIleGlnGlyGlnGluLeuS 134
|||||
463 TATCAAGGGCGCGGCTCAAGAGAGGAGGAGATTGAGGGCCAGGAGTTAT 512
134 eAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
|||||
513 CAGCCTACCAGGCTCCTCCCGACCGCACACAGTGGCTTCCATCCTCTAC 562
151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProly 167
|||||
563 CAGTCTTTGTCTATCTTTCAGGAAGGAAAGTCATCTCTCTCTCTCCCA 612
167 sGluAsnLysThrArgLysSerTrpLysMetAspArgPheLeuAsnArgp 184
|||||
613 GGAAACAAACCTCAGGCTCTTGGAAATGGACAGATTCTTGAACCGTT 662
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
663 TCCACCTGGCGCAACCTGAAAGCAGCCAGTTCATGACCCAGCAACTAC 712
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
|||||
713 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGGCCAGCGGCCCAA 762
217 sHisLysAsnGlnAlaGluLeuAlaLacys 227
|||||
763 GCACAAACACCGCGGAGATAGTCTGCCTGC 793

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT: AAA78426
seq_documentation_block:
```

ID AAA78426 standard; cDNA; 1028 BP.
AC AAA78426;
XX
XX 20-NOV-2000 (first entry)
XX
XX Human secreted protein gene 46 SEQ ID NO:56.
XX
XX Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW immune system disorder; angiogenesis; hyperproliferative disorder;
KW cardiovascular disorder; apoptosis; neurological disease;
KW infectious disease; wound healing; ss.
XX
XX Homo sapiens.
XX
XX WO200035937-A1.
XX
XX 22-JUN-2000.
XX
XX 16-DEC-1999; 99WO-US29950.
XX
XX 17-DEC-1998; 98US-0112809.
XX
XX 18-DEC-1998; 98US-0113006.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, NI J;
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
PI
XX WPI: 2000-431566/37.
DR
DR P-PSDB: AAB24482.
XX
XX Forty seven human nucleic acids encoding secreted proteins, useful in
PT the treatment, prevention and diagnosis of cancers, disorders of the
PT immune system, angiogenesis disorders, neurological diseases and
PT hyperproliferative disorders -
XX
XX Claim 1; Page 475; 562pp; English.
PS
XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the
CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; antianaemic;
CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
CC antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
CC nootropic; neuroprotective; antimicrobial and antiparkinsonian.
CC Human secreted protein polynucleotides, polypeptides, antagonists and/or
CC agonists may be useful in treating, preventing, and/or diagnosing other
CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
CC disorders of the immune system; (c) angiogenesis disorders; (d)
CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
CC associated with increase apoptosis; (g) neurological diseases; and
CC (h) infectious diseases. They are also used to promote wound healing.
CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 1028 BP; 309 A; 268 C; 259 G; 186 T; 6 other;
S0

alignment_scores:
Quality: 1218.00 Length: 227
Ratio: 5.389 Gaps: 0
Percent Similarity: 99.559 Percent Identity: 99.559
alignment_block:
US-09-215-435-225 x AAA78426 ..
Align seg 1/1 to: AAA78426 from: 1 to: 1028
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT: AAC77539
seq_documentation_block:
ID AAC77539 standard; cDNA; 909 BP.
XX
XX AAC77539;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF3094 polynucleotide sequence SEQ ID NO:6187.
DE
XX
XX Human; open reading frame; OREX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW

seq_name:	/cgml_9/gcgcdata/geneseq/NA2000.DAT:AAAL5582
seq_documentation_block:	
ID	AAAL5582 standard; cDNA; 903 BP.
XX	
XX	AAAL5582:
XX	
XX	01-AUG-2000 (first entry)
DT	
XX	
XX	Human phospholipid binding protein 2, PLBP2 gene.
XX	
XX	Human; phospholipid binding protein; PLBP2; foetal development
KW	reproduction disorder; cell proliferation disorder; immun-
KW	autoimmune disorder; AIDS; infertility; cytostatic; immuno-
KW	gene therapy; hereditary neuropathy;
KW	phosphatidylethanolamine binding protein D1; PE-BP D1; ss
XX	
XX	Homo sapiens.
OS	
XX	
XX	Key Location/Qualifiers
CD	88..771
TT	DSS

FT /*tag= a
XX /product= "Human PLBP2"
XX

US6063767-A.

16-MAY-2000.

09-DEC-1998; 98US-0208718.

28-OCT-1997; 97US-0958820.

(INCY-) INCYTE PHARM INC.

Corley NC, Shah P, Lal P, Hillman JL;

WPI: 2000-375529/32.

P-PSDB; AAY94263.

New purified phospholipid binding proteins 1 and 2 useful for
diagnosing, treating or preventing diseases disorders associated with
fetal development, reproduction, cell proliferation, and the immune
response

Example 5; Fig 2; 37pp; English.

The present sequence is the phospholipid binding protein 2 (PLBP2) gene.
This gene is expressed in lung, prostate and heart tissues. Also, the
protein is expressed in foetal tumour tissues. PLBP2 may be used for the
diagnosis, prevention, or treatment of disorders associated with foetal
development (e.g. hereditary neuropathies), reproduction (e.g.
infertility), cell proliferation (e.g. cancers), and the immune response
(AIDS). PLBP2 antibodies may also be developed for potential drug
screening or to quantitate PLBP2 gene expression in biopsied tissues. The
PLBP2 gene may be administered for gene therapy of disorders associated
with PLBP2. PLBP2 has high homology with the phosphatidylethanolamine
binding protein DI, PE-BP DI, of Onchocerca volvulus. PE-BP DI is thought
to play a role in transport or signal mechanisms between membranes and
the cytoplasm.

Sequence 903 BP; 222 A; 251 C; 242 G; 188 T; 0 other;

alignment_scores:
Quality: 1213.00 Length: 227
Ratio: 5.367 Gaps: 0
Percent Similarity: 99.559 Percent Identity: 99.119
alignment_block:
US-09-215-435-225 x AAA15582 ..
Align seg 1/1 to: AAA15582 from: 1 to: 903
1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
88 ATGGGTGGACAAATGAGGCTGGTGCACAGCAGCAGCTTTACTGGGTCTCAT 137
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
138 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 187
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
188 CCCTCTGGACGAGGACACCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 237
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
238 CCAGAGTTGGGAACATTGGCTGCAAGGTTGTCTCTGATTGTAACAAC 287
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
288 CAGACAGAAAGATCACCTCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 337
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100

|||||
338 CCGTGCACGCGCAACCTATATCCTGCTGATGCTGATCCAGATGCCCT 387
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
388 AGCAGAGCAGAACCCACAGACAGATTCTGGAGACATTGGCTGGTAACAGA 437
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
438 TATCAAGGGCGCGACCTGAAGGAAGGAGATTTCAGGCCAGGAGTTAT 487
134 exAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
488 CAGCCTATCCAGGCTCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 537
151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProLy 167
|||||
538 CAGTTCTTTGTCTATCTTCAGGAAGGAAAGTATCTCTCTCTCCCAA 587
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
588 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTTGAACCGCT 637
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
638 TCCACCTGGCGCAACCTGAAGCAAGCACCAGTTTCATGCCAGAACTAC 687
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
688 CAGGACTCACCACCTCCAGGCTCCAGAGGAAGGGCCAGCGAGGCCAA 737
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
738 GCACAAAAACCCAGCGGAGATAGTCCCTGCC 768
seq_name: /cgnl_9/gcdata/geneseq/geneseqn/NA2000.DAT:AAA78431
seq_documentation_block:
ID AAA78431 standard; cDNA; 952 BP.
XX
AC AAA78431;
XX
DT 20-NOV-2000 (first entry)
XX
DE Human secreted protein gene 46 SEQ ID NO:61.
XX
KW Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW immune system disorder; angiogenesis; hyperproliferative disorder;
KW cardiovascular disorder; apoptosis; neurological disease;
KW infectious disease; wound healing; ss.
XX
OS Homo sapiens.
XX
PN WO200035937-Al.
XX
PD 22-JUN-2000.
XX
PF 16-DEC-1999; 99WO-US29950.
XX
PR 17-DEC-1998; 98US-0112809.
XX
PR 18-DEC-1998; 98US-0113006.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX
DR WPI; 2000-431566/37.
DR P-PSDB; AAB24487.
XX

PT Forty seven human nucleic acids encoding secreted proteins, useful in
PT the treatment, prevention and diagnosis of cancers, disorders of the
PT immune system, angiogenesis disorders, neurological diseases and
PT hyperproliferative disorders -
XX
PS Claim 1: Page 478; 562pp; English.

XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the
CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; antianaemic;
CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
CC antithratic; antipsoriatic; antimicrobial and antiparkinsonian.
CC nootropic; neuroprotective; antineoplastic; antitumor; antidiabetic;
CC agonists may be useful in treating, preventing, and/or diagnosing other
CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
CC disorders of the immune system; (c) angiogenesis disorders; (d) diseases
CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
CC associated with increase apoptosis; (g) neurological diseases; and
CC (h) infectious diseases. They are also used to promote wound healing.
CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 952 BP; 310 A; 247 C; 223 G; 172 T; 0 other;

alignment_scores:
Quality: 1198.00 Length: 229
Ratio: 5.278 Gaps: 2
Percent Similarity: 99.127 Percent Identity: 99.127

alignment_block:
US-09-215-435-225 x AAA78431 ..
Align seg 1/1 to: AAA78431 from: 1 to: 952

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuLeu 17
63 ATGGGTGGCAATGAGCTGGTGCACAGCAGCACTGTACTGGGTCTCAT 112
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGlu 34
113 GATGGTGTCTACTGAGACGAGGATGAGAACAGCCGCTGTGCCCATGAG 162
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
163 CCCTCTGGACGAGGACACCTCTTTTCCAGGGCCCTTGAAGTTTCTTAC 212
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
213 CCAGAGTTGGGGAACATTTGGCTGCAAGGTTGTCTGATTCTAACAACTA 262
67 rArgGlnLysIleThrSerTrpMetGlu. ProIleValLysPheProGly 83
263 CAGACAGAAGATCACTCTCTGGATGGAGAGCCGATAGTCAAGTCCCGGG 312
84 AlaValAspGlyAlaThrTyr. IleLeuValMetValAspProAspAlap 100
313 GCGGTGGAGCGGCACTTAATCTCTGTTGATGATGATGATGATGATGAT 362
100 roSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThr 116
363 CTAGCAGACAGCAACCCACACACAGATTCTGGAGACATTGGCTGGTAACA 412
117 AspIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLe 133
413 GATATCAAGGGCGCCGACCTGAAGAAGGAAGATTCAGGGCCAGAGTT 462
133 uSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgT 150
463 ATCAGCTTACCAGGCTCCCTCCCGCCGACACACAGTGGCTTCCATCGCT 512

150 yrGlnPhePheValtyrLeuGlnGluGlyLysValIleSerLeuLeuPro 166
513 ACCAGTTCTTGTCTATCTTCAGAGAGAAAGTCATCTCTCTCTTCCC 562
167 LysGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnAr 183
563 AAGGAAACAAACATCGAGGCTCTTGGAAATGGACAGATTCTTGAACCG 612
183 gPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnT 200
613 TTTCCACCTGGCGGAACCTGAAGCAAGCACCAGCTTCATGACCCAGAAT 662
200 yrGlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluPro 216
663 ACCAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCCAGCGAGGCC 712
217 LysHisLysAsnGlnAlaGluIleAlaAlaCys 227
713 AAGCACAAACACAGCGGAGATAGCTGCCTGC 745

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA96350
seq_documentation_block:
ID AAA96350 standard; cDNA; 890 BP.
XX AC AAA96350;
XX DT 08-FEB-2001 (first entry)
XX DE cDNA encoding a novel polypeptide designated PRO4408.
XX KW Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;
KW PRO1889; PRO1890; PRO1887; PRO1785; PRO4353; PRO4357; PRO4405; PRO4356;
KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;
KW PRO4424; PRO4422; PRO4430; PRO4499; tumour; obesity; diabetes;
KW insulinemia; kidney disorder; Bergers disease; nephropathy;
KW Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis;
KW Crohns disease; ss.
XX OS Homo sapiens.
XX EH Key Location/Qualifiers
FT CDS 89..760
FT sig_peptide /*tag= a
FT /*tag= b
XX FT 89..154
XX PN WO200056889-A2.
XX PD 28-SEP-2000.
XX PF 01-MAR-2000; 2000WO-US05601.
XX PR 23-MAR-1999; 99US-0125774.
PR 23-MAR-1999; 99US-0125778.
PR 24-MAR-1999; 99US-0125826.
PR 31-MAR-1999; 99US-0127035.
PR 05-APR-1999; 99US-0127706.
PR 21-APR-1999; 99US-0130359.
PR 27-APR-1999; 99US-0131270.
PR 27-APR-1999; 99US-0131272.
PR 27-APR-1999; 99US-0131291.
PR 04-MAY-1999; 99US-0132371.
PR 04-MAY-1999; 99US-0132379.
PR 04-MAY-1999; 99US-0132383.
PR 25-MAY-1999; 99US-0135750.
PR 08-JUN-1999; 99US-0138166.
PR 20-JUL-1999; 99US-0144791.
PR 03-AUG-1999; 99US-0146970.
PR 09-DEC-1999; 99US-0170262.
XX PA (GETH) GENENTECH INC.
XX XX

PI Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2000-628263/60.
DR P-PSDB; AAB18923.
XX Novel secreted and transmembrane polypeptides useful for diagnosing
PT tumour in a mammal, for identifying agonists and antagonists of the
PT polypeptide and for therapeutic use
XX Claim 2; Fig 29; 222pp; English.
XX The present sequence encodes a secreted or transmembrane polypeptide.
XX The specification describes polypeptides designated PRO1484, PRO4334,
CC PRO1122, PRO1889, PRO1890, PRO1887, PRO1785, PRO4353, PRO4357, PRO4405,
CC PRO4356, PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,
CC PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is
CC useful for diagnosing tumour in a mammal. The polypeptides, their
CC agonists and antagonists are useful treating a condition associated with
CC expression or activity of the polypeptide. Conditions treated include
CC obesity, diabetes or hyper- or hypo-insulinemia. The polypeptides are
CC capable of inducing proliferation of mammalian kidney mesangial cells
CC and are therefore useful for treating kidney disorders associated with
CC decreased mesangial cell function such as Bergers disease or other
CC nephropathies associated with Schonlein-Henoch purpura, celiac disease,
CC dermatitis herpetiformis or Crohns disease. The nucleic acids may be used
CC to generate transgenic animals for use in development and screening of
CC therapeutically useful reagents and also for chromosome identification
CC and tissue typing.
XX Sequence 890 BP; 228 A; 246 C; 234 G; 182 T; 0 other;

alignment_scores: Quality: 1199.00 Length: 227
 Ratio: 5.261 Gaps: 0
Percent Similarity: 99.559 Percent Identity: 99.559
alignment_block:
US-09-215-435-225 x AAA96350 ..
Align seg 1/1 to: AAA96350 from: 1 to: 890
1 MetGlyTrpThrMetArgLeuValThrAlaLaLeuLeuLeuGlyLeuMe 17
89 ATGGGTGGACATGAGCTGGTGCACAGCAGCAGCTGTACTGGGTCTCAT 138
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
139 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 188
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
189 CCTCTTGGACGAGGACACCCCTCTTTTGCAGGGCTTGAAGTTTCTAC 238
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
239 CCAGAGTTGGGAACATGGCTGCAGAGGTGTCTCTGATTGTAACAACATA 288
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
289 CAGACAGAGATCACCTCCTCGATGGAGCCGATAGTCAAGTCCCGGGG 338
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
339 CCGTGGACGGCGCAACCTATATCTCTGGTATGGTGGATCCAGATGCCCT 388
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
389 ACAGAGCAGAACCCACAGACAGAGATCTCGAGACATGTGCTGGTACACA 438
117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuS 134

439 TATCAAGGGCGCCGACCTGAAGAAAGGGAAGATTTCAGGCCAGGAGTTAT 488
134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
489 CAGCCTACCAAGGCTCCCTCCCAACCGGCACACAGTGGCTTCCATCGCTAC 538
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProly 167
539 CAGTTCTTTGTCTATCTTCAGGAAGGAAAAGTCATCTCTCTCTCTCCCAA 588
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
589 GAAACAAACAAATCGAGGCTCTTGGAAAATGGACAGATTCTCAACCGCT 638
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
639 TCCACTGGGCGCAACCTGAAGCAAGCACCAGTTCATGACCAGCACTAC 688
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
689 CAGGACTCACCAACCTCCAGGCTCCACAGGAAGGCGCAGCGAGCCCAA 738
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
739 GCACAA.AACCCAGGCAGAGATAGTGCCTGC 768
seq_name: /cgcnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:AAx40582
seq_documentation_block:
ID_AAX40582 standard; cDNA: 447 BP.
XX AC AAX40582;
XX XX
XX 18-JUN-1999 (first entry)
XX Human secreted protein 5' EST SEQ ID No: 182.
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide; prostate;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX Homo sapiens.
XX OS
XX WO9906550-A2.
XX PN
XX 11-FEB-1999.
XX PD
XX 31-JUL-1998; 98WO-IB01232.
XX PF
XX 01-AUG-1997; 97US-0905144.
XX PR
XX (GEST) GENSET.
XX PA
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX WPI; 1999-153780/13.
XX P-PSDB; AAY11860.
XX New isolated prostate-derived nucleic acids - used to develop
XX products which may have cytokine, immune regulatory, haematopoiesis
XX regulating, anti-inflammatory or tumour inhibition activity
XX Claim 1; Page 285; 675pp; English.
XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins expressed in prostate, and encode the proteins
XX given in AAY11716 to AAY11993 respectively. The proteins given represent
XX the signal peptide and an N-terminal fragment of a secreted protein. The
XX nucleic acid sequences can be used for producing secreted human gene
XX products. They can also be used to develop products for diagnosis and

CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ Sequence 447 BP; 98 A; 110 C; 136 G; 99 T; 4 other;

alignment_scores:
Quality: 652.00 Length: 121
Ratio: 5.433 Gaps: 0
Percent Similarity: 99.174 Percent Identity: 99.174

alignment_block:
US-09-215-435-225 x AAX40582 ..
Align seg 1/1 to: AAX40582 from: 1 to: 447

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
85 ATGGTGGGACAAATGAGCTGGTCCACAGCAGCAGCTGTTACTGGGTCTCAT 134
17 tmetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
135 GATGGTGGTCACTGGAGCAGGAGGATGAGACAGCCCGTGTCGCCCATGAGG 184
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
185 CCCTCTGGACGAGACACCTCTTTTGCAGGGCCCTTGAAGTTTCTAC 234
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
235 CCAGAGTTGGGGAACATTTGGTGCAGAGTTGTTCCCTGATTGTDACAAC 284
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
285 CAGACAGAGATCACCCTCTCGATGGAGCCGATAGTCAAGTTCCCGGGGG 334
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
335 CCGTGGAGCGGGCAACCTATATCTCTGATGGTGATGGATGCCATGCCCT 384
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
385 ACCAGAGCAGAACCCAGACAGACAGATTCTGGAGACATTGGCTGGTAACAG 434
117 pileLysGlyAla 121
435 TATCAAGGGCGCC 447

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1992.DAT.AAQ30001
seq_documentation_block:
ID_AAQ30001 standard; cDNA to mRNA; 1047 BP.

XX AC AAQ30001;
XX 18-MAR-1993 (first entry)
DT HCNP precursor gene #1.
XX HCNP precursor gene #1.
XX Hippocampal cholinergic neurotrophic peptide; HCNP; hippocampal tissue;
KW neonatal rat; expression vector; neurodegenerative disorder; dementia;
KW Alzheimer's disease; Parkinson's disease; ss.
XX Rattus norvegicus.
OS

XX FH Location/Qualifiers
CDS 26..589
FT /*tag= a
FT 1..25
FT /*tag= b
FT 587..1047
FT /*tag= c
FT polyA_signal 987..992
FT /*tag= d
FT misc_RNA 29..61
FT /*tag= e
FT /label= Neurotrophic_activity_region
XX
XX EP511816-A.
XX 04-NOV-1992.
XX 27-APR-1992; 92EP-0303800.
XX 27-APR-1991; 91JP-0124688.
XX (OJIK/) OJIK K.
PA (SUMU) SUMITOMO PHARM CO LTD.
PA (YAMA/) YAMAMOTO M.
XX Aqul H, Fukushima N, Irie T, Kojima S, Nishihara T;
PI Ojika K, Ono K, Tohdoh N, Tojo S, Ueki Y;
XX WPI; 1992-367633/45.
DR P-PSDB; AAR27897.
XX Neurotrophic peptide derivs. - used for treating neurological
PT degenerative disorders, e.g. Alzheimer's disease or Parkinson's
PT disease
XX Claim 3; Page 31-32; 57pp; English.
XX The sequence given is the hippocampal cholinergic neurotrophic peptide
CC (HCNP) gene. The active part of the protein encoded by this sequence
CC is located in the first eleven amino acids at the N terminus. This
CC gene was obtained by preparing mRNA from the hippocampal tissue of
CC neonatal rats, 12 days after birth. The cDNA can be used in expression vectors to
CC produce both the full length protein and the active peptide fragment.
CC The gene and the peptide are useful for the treatment of
CC neurodegenerative disorders and dementia eg. Alzheimer's and
CC Parkinson's disease.
XX Sequence 1047 BP; 257 A; 296 C; 290 G; 204 T; 0 other;

alignment_scores:
Quality: 237.50 Length: 185
Ratio: 2.284 Gaps: 6
Percent Similarity: 56.216 Percent Identity: 32.973

alignment_block:
US-09-215-435-225 x AAQ30001 ..
Align seg 1/1 to: AAQ30001 from: 1 to: 1047
52 GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 68
131 GAGCTGGGCAAGTGTG.....CTGAGCCGCCACCCAGGTCATGAATAG 171
68 gGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAlav 85
172 ACCAAGCAGCATTTTCATGG.....GATGGCCTTG 200
85 alaSpGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 101
201 ATCTCTGGGAAGCTCTACACCCCTGCTCTCACAGACCCCGATGCTCCAGC 250

204ProThrLeuGlnAlaP 209
586 GTAGGGCGCTGAGAGCCGCGAGCCCGGGGACCCACAGTACAGTCAA 635
209 roArg 210
636 GTCGT 640

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1992.DAT:AAQ30002

seq_documentation_block:
ID AAQ30002 standard; cDNA to mRNA; 1447 BP.
AC AAQ30002;
XX 18-MAR-1993 (first entry)
XX HCNP precursor gene #2.
XX Hippocampal cholinergic neurotrophic peptide; HCNP; hippocampal tissue;
KW expression vector; neurodegenerative disorder; dementia;
KW Alzheimer's disease; Parkinson's disease; SS.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 120..680
FT CDS /*tag= a
FT 5'UTR 1..119
FT /*tag= b
FT 681..1447
FT /*tag= c
FT 123..155
FT /*tag= d
FT /*label= HCNP

XX EP511816-A.
XX 04-NOV-1992.
XX 27-APR-1992; 92EP-0303800.
XX 27-APR-1991; 91JP-0124688.
XX (OJIK/) OJIK K.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX (YAMA/) YAMAMOTO M.

XX Agui H, Fukushima N, Irie T, Kojima S, Nishihara T;
PI Ojika K, Ono K, Tohdoh N, Tojo S, Ueki Y,
XX WPI; 1992-367633/45.
XX P-PSDB; AAR27718.
XX Neurotrophic peptide derivs. - used for treating neurological
PT degenerative disorders, e.g. Alzheimer's disease or Parkinson's
PT disease
XX Disclosure; Page 42-44; 57pp; English.

XX The sequence given is the human hippocampal cholinergic neurotrophic
XX peptide (HCNP) precursor gene. The active part of the protein encoded
XX by this sequence is located in the first eleven amino acids at the N
XX terminus. This gene was isolated using the rat HCNP precursor gene
XX sequence as a guide (see also AAQ30001). From studies of the
XX structure of the rat HCNP precursor gene and protein the human HCNP
XX was determined. The rat and human HCNP's have markedly different
XX structures and composition. HCNP's are useful for the treatment of
XX neurodegenerative disorders and dementia eg. Alzheimer's and
XX Parkinson's disease.

XX Sequence 1447 BP; 342 A; 351 C; 406 G; 348 T; 0 other;

alignment_scores:
Quality: 208.50 Length: 206
Ratio: 1.813 Gaps: 7
Percent Similarity: 55.825 Percent Identity: 31.553
alignment_block:
US-09-215-435-225 x AAQ30002 ..
Align seg 1/1 to: AAQ30002 from: 1 to: 1447
3 TrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetv 19
74 TGGCTACCGGGGACTCCCGCTGCAGCGCTCTGTTGGCTCGCCATGC 123
19 alValThrGlyAspGluAsnSerProCysAlaHisGluAlaLeu 35
124 CGGTG.....GACCTCAGCAAGTGTCCGGCCCTTGAGCGTG 161
36 LeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr..... 50
162 CAAGAAGTGGAGCAGCAGCGCGCTGCAITGCTACCTAGCCGG 211
51ProGluLeuGlyAsnIleGlyCysLysValValProAspC 64
212 GCGCGGGTGGAGCAGCTGGGCAAGTG.....CTGACGCCACCC 252
64 ysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleValLys 80
253 AGGTTAAGAATAAGACCACCATTTCTGTGG..... 284
81 PheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValAspPr 97
285 ...GATGGTCTTGTATCAGGGAAGCTTACACCTTGGTCTGCACAGACC 331
97 oAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpL 114
332 GGATGCTCCAGCAGGAGGATCCCAATACAGAGATGGCATCTTTCC 381
114 euValThrAspIleLysGlyAlaAspLeuLysLysIleGlnGly 130
382 TGGTGTCAACATGAAGGGCAATGACATCAGCAGTGGCAGATC..... 425
131 GlnGluLeuSerAlaTyrGlnAlaProSerProAlaHisSerGlyPh 147
426CTCTCCGATTATGTGGGCTCGGGGCTCCCAAGGGCAGCGCT 469
147 eHisArgTyrGlnPhePheValThrLeuGlnGlyLysValIleSer. 163
470 CCACCGCTATGTCTGGCTGTTTACGAGCAGGACGCGCGCTAAAGTGTG 519
164LeuLeuProLysGluAsnLysThrArgGlySerTrpLys 176
520 ACAGAGCCCTCTCAGCAACCGATCTGGAGACCACCGTGGCAATTCAG 569
177 MetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerTh 193
570 GTGGCGTCTCTCCGTAAGAATATGAGTACGAGCGCGCGCTGGCTGGCAC 619
193 rGlnPheMetThrGln 198
620 GTGTTACCGGCGGCG 635

OM of: US-09-215-435-225 to: Issued_Patents_NA:* out_format : pfs
Date: Nov 15, 2001 4:23 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB=Issued_Patents_NA -CFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-215-435-225
Query length: 227
Database: Issued_Patents_NA*
Database sequences: 351203
Database length: 11323899
Search time (sec): 228.030000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4		1213.00	2632.22	5.8e-139	903
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-403-378B-3		212.50	436.77	1.1e-16	1047
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-403-378B-14		208.50	424.32	5.5e-16	1447
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-644-372-1		201.50	415.40	1.7e-15	822
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-1		164.00	332.73	7.0e-11	855
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-3		164.00	332.71	7.0e-11	855
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-998-416-1021		122.50	243.59	6.4e-06	726
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-998-416-991		115.00	227.69	4.9e-05	692
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2		113.50	124.78	26.68	4403765
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-1		113.50	124.76	26.75	4411529
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-47		87.50	167.26	0.1148	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-11		87.50	167.26	0.1148	700
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-082-614A-47		87.50	167.26	0.1148	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-8		87.50	145.38	1.90	4794
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-45		87.50	145.37	1.90	4795
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-082-614A-45		87.50	145.37	1.90	4795
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-46		82.50	156.30	0.4684	700
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-082-614A-46		82.50	156.30	0.4684	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-10		82.00	155.22	0.5380	699
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-07-765-830A-5		80.00	130.28	13.18	4258
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-288-830A-1		79.50	142.39	2.79	1333
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-785-241-2		79.00	131.95	10.63	3031
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-056-105-10		77.50	123.21	32.62	4895
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-056-105-9		77.50	113.50	113.34	11495
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-814-095-7		77.00	99.72	663.67	35060
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-697-610-1		76.00	128.82	15.90	2240
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-349-357-1		76.00	128.82	15.90	2240
/cgnl_7/ptodata/1/ina/PCPUS_COMB.seq:PCT-US95-16980-1		76.00	128.23	17.15	2359
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-211-417-2		76.00	124.46	27.81	3286
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-770-379-19		76.00	97.52	880.67	35100
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-757-669A-19		75.00	133.19	9.07	1257
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-791-115B-8		75.00	122.71	34.80	3160
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-791-115B-2		74.00	110.22	172.63	7812
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-09-368-590-1		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-812-13		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-812-13		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-921-177-13		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-362-577C-13		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-920-828-13		74.00	107.98	230.18	9515
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-492-027A-9		73.50	125.00	25.94	1935

/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-999-774A-5 - 73.50 120.27 47.57 2932
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-189-760-5 + 73.00 126.58 21.18 1529
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-188-811-5 + 73.00 126.58 21.18 1529
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-514-422-5 + 73.00 126.58 21.18 1529
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-656-984A-33 + 73.00 125.46 24.44 1687

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4

seq_documentation_block:
; Sequence 4, Application US/09208718
; Patent No. 6063767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PHOSPHOLIPID BINDING PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09208718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0379 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGTUT12
; CLONE: 3126479

alignment_scores:
Quality: 1213.00 Length: 227
Ratio: 5.367 Gaps: 0
Percent Similarity: 99.559 Percent Identity: 99.119

alignment_block:
US-09-215-435-225 x US-09-208-718-4 ...
Align seg 1/1 to: US-09-208-718-4 from: 1 to: 903
1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuLeuLeuLeu 17
|||||
88 ATGGGTGGACATGAGCTGGTGCACAGCAGCAGCTGTACTGGTCTCAT 137
17 tMetValValThrGlyAspGluAsnSerProCysAlaHisGluA 34
|||||
138 GATGGTGGTCACTGGAGACGAGGATGAGAACGCCGCTGTGCCCATGAGG 187

```

34 laLeuLeuaspGluaspThrLeupheCysGlnGlyLeuGluValPheTyr 50
|||||
188 CCCTCTGGACGAGGACACCCCTTTTGGCAGGGCTTGAAGTTTCTAC 237
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProaspCysAsnTyr 67
|||||
238 CCAGAGTGGGGAACATTGGCTGCAAGGTTGTTCTCGATTGTAACAATA 287
|||||
67 rArgGlnLysIleThrSerTrpMetLuprolleValLysPheProGlyA 84
|||||
288 CAGACAGACATCACCTCTGGATGGAGCGATAGTCAAGTTCCCGGGGG 337
|||||
84 laValaspGlyAlaThrTyrIleLeuValMetValaspProaspAlaPro 100
|||||
338 CCGTGGAGCGGCAACCTATATCTGCTGGTATGTTGATCCAGATGCCCT 387
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
388 AGCAGACAGACACCCAGACAGATTCTGGAGACATTGGCTGGTAACAGA 437
|||||
117 PileGlyGlyAlaaspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
438 TATCAAGGGCGCCGACCTGAAGGAAGGAAGATTCAGGGCCAGAGTTAT 487
|||||
134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
488 CAGCTACCAGGCTCCCTCCCGCCGACACACAGTGGCTTCCATCGCTAC 537
|||||
151 GlnPheheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
|||||
538 CAGTTCTTGTCTATCTCAGGAAGAAAGTATCTCTCTCTTCCCA 587
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
|||||
588 GGAACAACTCGAGGCTCTTGGAAATGGACAGATTCTGAAACGCT 637
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
638 TCCACCTGGGGAACTGAACAAAGACCCAGTTTCATGACCCAGAACTAC 687
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
|||||
688 CAGGACTCACCACCTCCAGGCTCCAGAGGAAGGGCCAGGAGCCCAA 737
|||||
217 shLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
738 GCACAAAAACCGGCGAGATAGTGCCTGC 768
|||||

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:us-08-403-378B-3

seq_documentation_block:
; Sequence 3, Application US/08403378B
; Patent No. 575991
; GENERAL INFORMATION:
; APPLICANT: TOHDOH, NAOKI
; APPLICANT: TOJO, SHIN-ICHIRO
; APPLICANT: KOJIMA, SHIN-ICHI
; APPLICANT: UKI, YASUYUKI
; APPLICANT: NISHIHARA, TOSHIO
; APPLICANT: FUKUSHIMA, NOBUYUKI
; APPLICANT: IRIE, TSUNEMASA
; APPLICANT: ONO, KEIICHI
; APPLICANT: AGUI, HIDEO
; APPLICANT: OJIKI, KOSHI
; TITLE OF INVENTION: NEUROTROPHIC PEPTIDE DERIVATIVES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.

```

```

ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,378B
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-124688
FILING DATE: 27-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1-080398
FILING DATE: 30-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1-280590
FILING DATE: 27-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1-333241
FILING DATE: 21-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-243003
FILING DATE: 12-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/758,043
FILING DATE: 12-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/873,764
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/01214
FILING DATE: 27-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: BIGGART, WADDELL A
REGISTRATION NUMBER: 24,861
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1047 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rattus norvegicus
STRAIN: Wistar
TISSUE TYPE: hippocampal tissue of brain
IMMEDIATE SOURCE:
CLONE: A010-12
NAME/KEY: CDS
LOCATION: 26..586
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..25
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 26..586
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 587..1047
FEATURE:
NAME/KEY: polyA_site
LOCATION: 1008..1047
FEATURE:

```


Thu Nov 15 10:52:09 2001

```
;
;
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..119
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 120..680
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 681..1447
;
; US-08-403-378B-14
;
alignment_scores:
  Quality: 208.50      Length: 206
  Ratio: 1.813        Gaps: 7
Percent Similarity: 55.825 Percent Identity: 31.553

alignment_block:
US-09-215-435-225 x US-08-403-378B-14 ..
Align seg 1/1 to: US-08-403-378B-14 from: 1 to: 1447
3 TrpThrWetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetV 19
|||
|||
74 TGGCCTACCGCGCAGCTCCCGGCTGGACGCTCTGCTTGGCTCGCCATGC 123
|||
|||
19 alValThrGlyAspGluAspGluAsnSerProCysAlaHisGluAlaLeu 35
|||
|||
124 CGGTG.....GACCTCAGCAAGTGGTCCGGGCCCTTGAGCCTG 161
|||
36 LeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr..... 50
|||
162 CAAGAAGTGGCAGCAGCAGCGCGCAGCCCGCTGCATGTCACCTACGCCGG 211
|||
51 .....ProGluLeuGlyAsnIleGlyCysLysValValProAspC 64
|||
212 GCGCGCGTGGACGAGCTGGGCAAGTG.....CTGACGCCACCC 252
|||
64 ysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleValLys 80
|||
253 AGGTTAAGATAGACCCACCAAGCATTCGTGG..... 284
|||
81 PheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValAspPr 97
|||
285 ...GATGCTCTTGATTCAGGAAGCTCTACACCTGGTCTCGACAGACC 331
|||
97 oAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpL 114
|||
332 GGATGCTCCAGCAGGAGGATCCCAATACAGAGAAATGGCATCATTTCC 381
|||
114 euValThrAspIleGlyAlaAspLeuLysGlyLysIleGlnGly 130
|||
382 TGGTGGTCAACATGAAGGGCAATGACATCAGCAGTGGCAGAGTC..... 425
|||
131 GlnGluLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPh 147
|||
426 .....CTCTCGATATATGCGCTCGGGCTCCCAAGGCACAGGCCT 469
|||
147 eHisArgTyrGlnPhePheValTyrLeuGlnGlyLysValIleSer. 163
|||
470 CCACCGCTATGCTGGTGGTTTACGACAGCAGCAGCGCGCTAAAGTGTG 519
|||
164 .....LeuLeuProLysGluAsnLysThrArgGlySerTrpLys 176
|||
520 ACAGGCCATCTCCAGCAACCCATCTGGAGACCCCGTGGCAAAATTCAG 569
|||
177 MetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerTh 193
|||
570 GTGGCGTCTCTCGTAAAGATGATGAGTCTCAGGCGCCCGGTGGTGGCAC 619
|||
193 rGlnPheMetThrGln 198
|||
620 GTCTTACCAGGCGCAG 635
|||
```

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-644-372-1

seq_documentation_block:

Sequence 1, Application US/07644372

Patent No. 5416009

GENERAL INFORMATION:

APPLICANT: Lazzeri, Mario E.

APPLICANT: Nutman, Thomas B.

APPLICANT: Weiss, Niklaus

TITLE OF INVENTION: A DNA SEGMENT ENCODING A SPECIFIC

TITLE OF INVENTION: IMMUNODIAGNOSTIC ANTIGEN

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

STREET: 1615 L. Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/644,372

FILING DATE: 19910123

CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)861-3000

TELEFAX: (202)822-0944

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 822 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: Double

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 52..507

US-07-644-372-1

alignment_scores:

Quality: 201.50 Length: 147

Ratio: 2.144 Gaps: 7

Percent Similarity: 63.946 Percent Identity: 36.054

alignment_block:

US-09-215-435-225 x US-07-644-372-1 ..

Align seg 1/1 to: US-07-644-372-1 from: 1 to: 822

55 AsnIleGlyCysLysValValPro...AspCysAsnAsnTyrArgGlnLy 70

|||||

244 AATCTGGCAATGAACCTTACGCCGAGGTAAGAATCAGCCGACAAA 293

|||||

70 sIleThrSerTrpMetGluProIleValLysPheProGlyAlaValAspG 87

|||||

294 AGTA...TCATGG.....GATCGGGAACCTG 316

|||||

87 lYalaThrTyrIleLeuValMetValAspProAspAlaProSerArgAla 103

|||||

317 GAGCCTTATATACGCTCGTTATGACTGATCCGACGCCACCATCTCGAAA 366

|||||

104 GluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysG 120

|||||

367 AACCCCGTATTACAGAGAGTGGCACCATTGTTGATAATTAATATTCTG 416

|||||

120 yAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyr. 136

|||||

417 ACAAATGTTAGCAGTGGCACAGTG.....TTATCTGATTATT 454

```

471  ..GGGTGGCGCCAGAACTTCAACACTGCGGAGTTTGCTGAGATCTACAAT 518
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
186  LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
      ||||| ||| |||: : : : : : : : : : : : : : : : :
519  CTCGGCCTTCCTGGCGCGCAGTTTCTACAAATGTCAG 557
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

seq_name: /cgnl_7/ptodata/1/lna/6B_COMB.seq:US-09-060-726A-3

seq_documentation_block:
; Sequence 3, Application US/09060726A
; Patent No. 6225530
; GENERAL INFORMATION:
; APPLICANT: Weigel, Detlef
; APPLICANT: Salk Institute
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; MODIFIED PLANTS HAVING MODULATED FLOWER
; FILE REFERENCE: SALKINS 026A
; CURRENT APPLICATION NUMBER: US/09/060,726A
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-060-726A-3

alignment_scores:
    Quality: 164.00      Length: 113
    Ratio: 2.412        Gaps: 4
    Percent Similarity: 60.177      Percent Identity: 36.283

alignment_block:
US-09-215-435-225 x US-09-060-726A-3/reverse ..

Align seq 1/1 to reverse of: US-09-060-726A-3 from: 1 to: 856

```

```

002  TATACCTTTGGTATGAGTGGGATCCAGATGTTCCAGAGTCTACAGAACCCCTCA 353
106  gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 123
    :||:  ::  ||||||||| ||||| |||||
552  CCTCGAGAAATATCTCCATTGGTGGTGAATGATATCCCTGCTACA... 507
123  euLysLysGlyLysIleGlnGlnGlnLeuSerAlaTyrGlnAlaPro 139
    ||||  ||||| :|||: |||: |||: |||: |||: |||: |||: |||:
506  ....ACTGGACAACCTTTGGCAATGAGATGTGTGTACGAAATCCA 462
140  SerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
461  ACTCCCACTGCA.....GGAATTCATCGTCTCGTGTATATTGTTTCG 418
156  uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 173
    ||||  ||||: ||||
417  ACAGCTTGGCAGGCAACACAGTGTATGCACCA..... 387
173  LysSerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 185
    :|||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
386  ..GGTGGGGCCAGACTTCAACACTCGCGAGTTTGCTGAGATCTACAA 339
186  LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
    ||||| |||| ||||| ||||| ||||| ||||| ||||| |||||
338  CTCGGCCTTCCCGTGGCGCAGTTTCTCAATTTGTCAG 300

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq;US-08-998-416-1021
seq_documentation_block:
; Sequence 1021, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:

```

Thu Nov 15 10:52:09 2001

us-09-215-435-225.rni

```

; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1021:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1025UP
; US-08-998-416-1021

```

```

alignment_scores:
  Quality: 122.50      Length: 89
  Ratio: 2.149         Gaps: 3
  Percent Similarity: 64.045  Percent Identity: 34.831

```

alignment_block:

US-09-215-435-225 x US-08-998-416-1021/rev ..

Align seg 1/1 to reverse of: US-08-998-416-1021 from: 1 to: 726

```

87 GlyAlaThrTrpIleLeuValMetValaspProAspAlaProSerArgAl 103
||| :::: ||| ||::||| ||::||| ||::||| ||::||| ||::|||
717 GGGGACCTATTCCGCTGGGATGACAGACCCAGATGCTCCCTCGCGGTC 668

103 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLys. 119
:::: :::: :::: ||::||| ||::||| ||::|||
667 GGACCACAAAGTGTGCGAATACTGCCACTTCTGGAAACGAACATAACGC 618

120 ..GlyAlaAsp.....LeuLysLysGlyLysIleGlnGlnGluLeu 133
||::||| :::: :::: :::: :::: ::::
617 TGGGCTCGGATGACGGGGGTGTCGACGTCGTCTAAAGGCGACCCCGCAG 568

134 SerAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
:::: :::: ||::||| ||::||| ||::||| ||::|||
567 GTGAGACACATGGGCCCTGCGCGCGCGCGGCACACACAGGGGCTCACGGTA 518

```

```

150 rGlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProL 167
| ::::||| ||| ||::||| :::: |||
517 CGTGTGTTCTTCTTTCGCGAGCCCGCGGCTA.....GAACTGAGCG 474

167 ySGluAsnLysThrArg 172
::||::||| |||
473 AGGAGGAGGTCACCCGC 457

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:us-08-998-416-991

seq_documentation_block:
; Sequence 991, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 991:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 692 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1609RP
; US-08-998-416-991

```

```

alignment_scores:
  Quality: 115.00      Length: 140
  Ratio: 1.620         Gaps: 8
  Percent Similarity: 50.714  Percent Identity: 30.000

```

alignment_block:

US-09-215-435-225 x US-08-998-416-991 ..

Align seg 1/1 to: US-08-998-416-991 from: 1 to: 692

26 GluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAspThrLeuP 42
|||||:|||||:|||||
336 GAAACGGTCCATCAGGACAC..... 356
42 ecysGlnGlyLeuGluValPheTyrPro.....GluL 53
||| ||| |||||
357CTGTAGTGGAGTACCCGGCGGAGTCTACAGCGGTGACGC 396
53 euGlyAsnIle.....GlyCysLysValValProAspCysAsn 65
|||||:|||||:|||||:|||||:|||||:|||||
397 TGGGGACGTTATGCTGTGGAGGCTACGACGCGTGCACAC..... 440
66 AsnTyrArgGlnLysIleThrSerTrpMetGluProIleValLysPhePr 82
|||:|||||:|||||:|||||:|||||:|||||
441CTGATGTTAATCAGCAGCCGAGCC 463
82 oglyAlaVal.....AspGlyAlaThrTyrIleLeuValMetValAspProA 98
|||||:|||||:|||||:|||||:|||||:|||||
464 GGAATCGTCAGGAGGGGAGGACCTATTACGCTGGCGATGACAGACCCAG 513
98 sPalapProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeu 114
|||||:|||||:|||||:|||||:|||||:|||||
514 ATGCTCCCTCGGGTGGCGGACCAAGTGTGCGAATACGACACCTTCTG 563
115 ValThrAspIleLys.....GlyAlaAsp.....LeuLysLysGlyLysI 128
|||||:|||||:|||||:|||||:|||||:|||||
564 GAAACGAACATACGCTGGGCTCGGATGACGGGTGTCGCACGTGTGTCT 613
128 eGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProProAlaHis 144
|||||:|||||:|||||:|||||:|||||:|||||
614 AAGGGCACCCGCGAGTGGAGCACATGCGCCCTGCGCGCGGGCGCGC 663
145 SerGlyPheHisArgTyr 150
:|||||:|||||:|||||:|||||:|||||:|||||
664 ACAGGGGCTCACCGGTAC 681

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2

seq_documentation_block:
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

alignment_scores:
Quality: 113.50 Length: 179
Ratio: 1.419 Gaps: 11
Percent Similarity: 44.693 Percent Identity: 27.374

alignment_block:

US-09-215-435-225 x US-09-103-840A-2/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-2 from: 1 to: 4403765

10 AlaAlaLeuLeuLeuGlyLeuMetMetValVal..... 20
||| :|||:|||||:|||||
2154533GCACTGGCCATCTTGGGCTGGCGCTCCCGTGGCGCTAGTTCCTACGG 2154484
21ThrGlyAspGluAspGluAsnSerProCysAlaHisGluAla.... 34
:|||||:|||||:|||||:|||||:|||||:|||||
2154483TGCAACAGGTGACAGTCGAAAGGCGCGCGCTGGCGCGGAAACAGCAG 2154434
35LeuLeuAsp 37
2154433CGCTCGGTGGAGTATGCCGAAACGCCCTACCGCGGATGTACTGACAATC 2154384
38 GluAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuG 54
:|||||:|||||:|||||:|||||:|||||:|||||
2154383AGCAGTCCGGCATTCGCCGACGCGTGGCGCATC.....CCGGAACAGTA 2154340
54 yAsnIle....GlyCysLysValValProAspCysAsnAsnTyrArgGlnL 70
:|||||:|||||:|||||:|||||:|||||:|||||
2154339CACCTGCAAGAGGAGCAATATCGCGCTCG..... 2154309
70 ystIleThrSerTrpMetGluProIleValLysPheProGlyAlaValAsp 86
:|||||:|||||:|||||:|||||:|||||:|||||
2154308.....TTGACCTGGTGGCGCG.....TTGGCGCGCA..... 2154279
87 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 103
:|||||:|||||:|||||:|||||:|||||:|||||
2154278.....CTCGTTGTCGATGATCCGCGACGCACCT.....CG 2154250
103 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLys. 119
:|||||:|||||:|||||:|||||:|||||:|||||
2154249CGAACCT.....TACGTCCATTGGATCGTGTGATCGGATCGCC 2154212
120GlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSer 134
:|||||:|||||:|||||:|||||:|||||:|||||
2154211CTGGTCTGGCAGCACCGCGGATGGTGAGACTCCCGTGGCGGAATCAGC 2154162
135AlaTyrGlnAlaProSerProAla 143
:|||||:|||||:|||||:|||||:|||||:|||||
2154161CTGCCGAACCTCCAGCGGTGACGCCGATACACCGGCCCTGCGCGCGGC 2154112
143 aHisSerGlyPheHisArgTyrGlnPhePheValTyr 155
:|||||:|||||:|||||:|||||:|||||:|||||
2154111GGCACCGGACACACACACTACCGGTTTACCCCTCTAC 2154075

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-1

seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

alignment_scores:

Quality: 113.50
Ratio: 1.419

Length: 179
Gaps: 11

Thu Nov 15 10:52:09 2001

us-09-215-435-225.rni

```

Percent Similarity: 44.693 Percent Identity: 27.374
alignment_block:
US-09-215-435-225 x US-09-103-840A-1 rev ..
Align seg 1/1 to reverse of: US-09-103-840A-1 from: 1 to: 4411529
10 AlaAlaLeuLeuGluCysGlnGlyMetValVal..... 20
2157273GCGACTGCGGCTGGGGCTGGCGTCCCGTAGTGGCTACGG 2157224
21 .....ThrGlyAspGluAspLeuSerProCysAlaHisGluAla.... 34
2157223TGGCAACGGTGACAGTGACAAAGGCGCGCGCTGGCGCGAAGCAG 2157174
35 .....LeuLeuAsp 37
2157173GCGTCGCGTAGTATGCCGGAACCCCTACCGCGGATGTACTGACAATC 2157124
38 GluAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuG1 54
2157123AGCAGTCGCGCATTCGCGGAGCGTGGCGCGATC.....CCGGAACAGTA 2157080
54 yAsnIle...GlyCysLysValValProAspCysAsnAsnTyrArgGlnL 70
2157079CACCTGCAAGAGGAGCAATATCGCGCTCGG..... 2157049
70 ySileThrSerTrpMetGluProIleValLysPheProGlyAlaValAsp 86
2157048.....TTGACCTGGTCGCGCGG.....TTTGGCGGCGCA..... 2157019
87 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 103
2157018.....CTCGTTGTGATGATCGCGGACGACCT.....CG 2156990
103 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLys. 119
2156989CGAACCT.....TACGTCCATTGGATCGGATCGCC 2156952
120 .....GlyAlaAspLeuLysGlyLysIleGlnGlyGlnLeuSer 134
2156951CTGGTGTGGCAGCAGCCGCGATGGTGGAGACTCCCGGTGGCGGAATCAGC 2156902
135 .....AlaTyrGlnAlaProSerProAl 143
2156901CTGCCGAACTCAGCGGTACCCCGCATACCGCGCCCTGCCGCGCGC 2156852
143 ahisSerGlyPheHisArgTyrGlnPhePheValTyr 155
2156851GGGCACGGGACACACCACTACCGGTTTACCCCTCTAC 2156815
seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-47

seq_documentation_block:
; Sequence 47, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.

```

```

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,185
FILING DATE: 12-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02356.0068-00000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-313-185-47

alignment_scores:
Quality: 87.50 Length: 108
Ratio: 1.620 Gaps: 7
Percent Similarity: 50.000 Percent Identity: 31.481

alignment_block:
US-09-215-435-225 x US-08-313-185-47 ..
Align seg 1/1 to: US-08-313-185-47 from: 1 to: 700
64 CysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleVal... 79
321 TGCATTTCAGCTGCAAGGGGGGCAACG...TGGCGCGCCACCGTTGACGTG 367
80 LysPheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValA 96
368 GTCTGTGGCGCGCGGCGGAGCAACTGGCA.....CTCGTCGTGATG 408
96 spProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 112
409 ACCCGGACGCGGTGGC.....GGACTGTACGTGCAC 440
113 TrpLeuValThrAspIle.....LysGlyAlaAsp..... 122
441 TGGATCGTGACCGGAATCCCGCTGCTGCGCAGCAGCGCGGATGTCA 490
123 .....LeuLysLysGlyLysIleGlnGlyG 131
491 GACTCTGTGGTGGGACAGCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
131 LngLysLeuSerAlaTyrGlnAlaProSerProAlaHisSerGlyPhe 147
540 .....TACTTGGTTCATGCCCGCGCGCGCGGCGGCGGCGGCGGCA 575
148 HisArgTyrGlnPhePheValTyr 155
576 CACCACTACCGGTTTACCCCTCTAC 599
seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-11

seq_documentation_block:
; Sequence 11, Application US/08459499
; Patent No. 5871912
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate

```

APPLICANT: Cole, Stewart T.
APPLICANT: Young, Douglas B.
APPLICANT: Zhang, Ying
TITLE OF INVENTION: Nucleic Acid Probes, Sequences, and Methods
TITLE OF INVENTION: for Detecting Mycobacterium Tuberculosis Resistant to Isoniazid
TITLE OF INVENTION: Amended
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,499
FILING DATE: 02-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/875,940
FILING DATE: 30-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/929,206
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/029,655
FILING DATE: 11-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0110-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-459-499-11

alignment_scores:
Quality: 87.50 Length: 108
Ratio: 1.620 Gaps: 7
Percent Similarity: 50.000 Percent Identity: 31.481
alignment_block:
US-09-215-435-225 x US-08-459-499-11 ..
Align seg 1/1 to: US-08-459-499-11 from: 1 to: 700
64 CysAsnAspTyrArgGlnLysIleThrSerTrpMetGluProIleVal... 79
|||||: : : : :
321 TGCATTCAGCTCAGGAGGGGCAAG...TGGCCGCCACCGTTGACGTG 367
|||||: : : : :
80 LysPheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetVal 96
: : : : :
368 GTCGTCGCCCGGGCGAGCAAGCA...CTCGTCGTGCATG 408
96 sProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 112
|||||: : : : :
409 ACCCGACGCGTCGCGC...GGACTGACGTGCAC 440
113 TrpLeuValThrAspIle.....LysGlyAlaAsp..... 122

|||||: : : : :
441 TGGATCGTCACCGGAATCGCCCTGGCTCTGGCAGCACGCGGATGGTCA 490
123 : : : : :
491 GACTCCTGCTGGTGGGCACAGCGTGGCCGAATTCGTGTGTCGGCAAGGA. 539
131 InGluLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPhe 147
|||||: : : : :
540 : : : : :
148 HisArgTyrGlnPheValTyr 155
|||||: : : : :
576 CACCACCTACCGGTTTACCCTCTAC 599
seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-082-614A-47
seq_documentation_block:
Sequence 47, Application US/09082614A
Patent No. 6124098
GENERAL INFORMATION:
APPLICANT: Heym, Beate
APPLICANT: Cole, Stewart
APPLICANT: Young, Douglas
APPLICANT: Zhang, Ying
APPLICANT: Honore, Nadine
APPLICANT: Telenti, Amelio
APPLICANT: Bodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
TITLE OF INVENTION: in Mycobacterium Tuberculosis
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,614A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,185
FILING DATE: 12-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02356.0068-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-082-614A-47

alignment_scores:
Quality: 87.50 Length: 108
Ratio: 1.620 Gaps: 7
Percent Similarity: 50.000 Percent Identity: 31.481
US-09-082-614A-47

THIS PAGE BLANK (USPTO)


```

; OTHER INFORMATION: seq AALLGLMMVVTG/DE
US-09-471-276-7

alignment_scores:
  Quality: 1224.00      Length: 227
  Ratio: 5.392         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-09-471-276-7

Align seg 1/1 to: US-09-471-276-7 from: 1 to: 826

1 MetGlyTTPThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
15 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 64
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
65 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCGCTGCCCATGAGG 114
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
115 CCCTCTCTGACGAGGACACCCCTCTTTTCCAGGGCCCTTGAAGTTTCTAC 164
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
165 CCAGAGTTGGGAACATTTGGCTGCAAGGTTTCTCTGATTGTAACAAC 214
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
215 CAGACAGAAAGATCACCTCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 264
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
265 CCGTGGACGGCGCAACCTATATCTCTGGTGATGGTGATCCAGATGCCCT 314
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
315 AGCAGAGCAGAACCCAGACAGACAGATTTCTGGAGACATTTGGCTGTAACA 364
117 pIleLysGlyAlaAspLeuLysLysLysLysLysLysLysLysLysLys 134
365 TATCAAGGGCGCGCACTGAAGAAGGGAAGATTCAGGGCCAGGAGTTAT 414
134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
415 CAGCCTACCGAGGCTCCCTCCCGCCGACACAGTGGCTTCCATCGCTAC 464
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
465 CAGTTCTTTGTCTATCTTTCAGGAAGGAAAGGTCATCTCTCTCTCCCAA 514
167 sGluAsnLysThrArgLysSerTrpLysMetAspArgPheLeuAsnArgP 184
515 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTTCTGAACCGTT 564
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
565 TCCACCTGGGCGAACCTGAAGCAAGCAGCCAGTTCATGACCCAGAACTAC 614
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGGCCAGGAGGCCCAA 664
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
665 GCACAAAACCCAGGCGGAGATAGTGCCTGC 695

seq_name: /cgnl_7/ptodata/1/pna/us094_COMB.seq:US-09-834-366-7
seq_documentation_block:
;
; Sequence 7, Application US/09471276
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: GENSET.025CP1
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCI/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: score 8.5

```



```
; Sequence 7, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVTG/DE
US-09-834-366-7

alignment_scores:
  Quality: 1224.00      Length: 227
  Ratio: 5.392          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-09-834-366-7 ..

Align seg 1/1 to: US-09-834-366-7 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTGGACAATGAGGTGGTCACAGAGCAGCTGTTACTGGTCTCAT 64
|||||
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGGTGGTCACTGGAGCAGAGATGAGAACAGCCCGTGTGCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCCTCTCGACGAGGACACCCCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 164
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
165 CCAGAGTTGGGGAACATTGGCTGCAAGGTGTCTCTGATTGTAACAACTA 214
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
215 CAGACAGAGATCACCCTCTCGATGAGCGCGATAGCAAGTTCCCGGGG 264
|||||
84 laValAspGlyAlaThrTrpIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCGTGGACGGCGCAACCTATATCTCGTGTGATGGTGCATCCAGATGCCCT 314
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
315 AGCAGAGACAGAACCCACAGACAGAGATTCTGGAGACATTGGCTGTAACAGA 364
|||||
117 pileLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
365 TATCAAGGGCGCGGCTCTGAAGAAAGGAAGATTTCAGGGCCAGAGTTAT 414
|||||
134 eAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
```

```
|||||
415 CAGCCTACCAGGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 464
|||||
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
|||||
465 CAGTTCTTTGTCTATCTTCTCAGGAAGAAAGGTCTCTCTCTCTCCCAA 514
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
515 GGAAAAACAAACTCGAGGCTCTTGAAAAATGGACAGATTCTGAACCGTT 564
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCCACCTGGCGCAACCTGAAGCAAGCACCAGTTCATGACCCAGAACTAC 614
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
|||||
615 CAGGACTCACCACCTCCAGGCTCCAGAGAAAGGGCCAGGAGCCCAA.664
|||||
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
665 GCACAAAAACACGCGGAGATAGTCCCTGC 695

seq_name: /cgnl_7/ptodata/1/pna/US6008_COMB.seq:US-60-081-563-70

seq_documentation_block:
; Sequence 70, Application US/60081563
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SECRETED PROTEINS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/081,563
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.027PR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Prostate
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 8.5
```

Thu Nov 15 10:52:10 2001

us-09-215-435-225.rnmpm

```

OTHER INFORMATION: seq AALLGLMMVVTG/DE
FEATURE:
NAME/KEY: poly_a_signal
LOCATION: 795..800
FEATURE:
NAME/KEY: poly_a
LOCATION: 814..826
IDENTIFICATION METHOD: blastn
FEATURE:
NAME/KEY: est
LOCATION: 426..664
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 99
OTHER INFORMATION: region 99..337
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 327..435
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 1..109
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 716..777
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 100
OTHER INFORMATION: region 391..452
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 660..720
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 334..394
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 506..630
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 90
OTHER INFORMATION: region 327..451
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 323..435
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 94
OTHER INFORMATION: region 143..255
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 180..280
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 2..102
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 464..528
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 284..348
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 279..321
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 100
OTHER INFORMATION: region 100..142
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est

```

```

LOCATION: 426..466
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 97
OTHER INFORMATION: region 245..285
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 358..674
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 141..457
OTHER INFORMATION: id AAL192427
FEATURE:
NAME/KEY: est
LOCATION: 670..779
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 37..146
OTHER INFORMATION: id AAL192427
FEATURE:
NAME/KEY: est
LOCATION: 485..602
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 224..341
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 601..709
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 91
OTHER INFORMATION: region 115..223
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 704..779
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 46..121
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 392..436
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 97
OTHER INFORMATION: region 393..437
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 429..465
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 94
OTHER INFORMATION: region 363..399
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 464..493
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 334..363
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 343..730
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 8..395
OTHER INFORMATION: id HSPD04880
FEATURE:
NAME/KEY: est
LOCATION: 507..663
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96

```

seq_documentation_block:
; Sequence 7, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-60-081-563-70 ..

Align seg 1/1 to: US-60-081-563-70 from: 1 to: 826

```
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeu 17
|||||
15 ATGGGTGGACATGAGGTGGTGCAGCAGCAGCTGTACTGGGTCTCAT 64
|||||
17 tMetValThrGlyAspGluAsnSerProCysAlaHisGlu 34
|||||
65 GATGGTGGTCACTGGAGACGAGGATGAGACGCCGCTGTGCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCCTCTGGACGAGGACACCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 164
|||||
51 ProGluLeuGlyAsnLeuGlyCysLysValValProAspCysAsnAsnTyr 67
|||||
165 CCAGAGTTGGGGAACATGGCTGCAAGGTGTTCCTGATTGTAACAACTA 214
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGly 84
|||||
215 CAGACAGAAGATCACCTCTGGATGGAGCGGATGTCAGATTCCCGGGG 264
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCCTGGACGGCGCAACTATATCTGGTGTGGTGCAGATGCCCT 314
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThr 117
|||||
315 AGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGTTAAACA 214
|||||
117 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGly 84
|||||
365 TATCAGAGCGCGCACTGAGAAAGGAGATTCAGGCCAGGAGTTAT 414
|||||
134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
415 CAGCCTACCAGGCTCCCTCCCGACCGCACACAGTGGCTTCCATCGCTAC 464
|||||
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuPro 167
|||||
465 CAGTCTTTGTCTATCTTCAGGAAGAAAGGTATCTCTCTCTCCCA 514
|||||
167 sGluAsnLysThrArgLysSerTrpLysMetAspArgPheLeuAsnArg 184
|||||
515 GGAACAAACAACTCGAGGCTCTTGGAAATGACACATTTCTGAACCGT 564
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCCACCTGGCGCAACTGAGCAAGCACCCAGTTTCATGCCAGAACTAC 614
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluPro 217
|||||
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGGCCAGCGAGCCAA 664
|||||
217 sHisLysAsnGlnAlaGluIleAlaCys 227
|||||
665 GCACAAACACCGCGGAGATAGCTGCCTGC 695
```

seq_name: /cgn1_7/ptodata/1/pna/US6014_COMB.seq:US-60-147-499-7

seq_documentation_block:
; Sequence 7, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVTG/DE
US-60-147-499-7

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-60-147-499-7 ..

Align seg 1/1 to: US-60-147-499-7 from: 1 to: 826

```
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeu 17
|||||
15 ATGGGTGGACATGAGGTGGTGCAGCAGCAGCTGTACTGGGTCTCAT 64
|||||
17 tMetValThrGlyAspGluAsnSerProCysAlaHisGlu 34
|||||
65 GATGGTGGTCACTGGAGACGAGGATGAGACGCCGCTGTGCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCCTCTGGACGAGGACACCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 164
|||||
51 ProGluLeuGlyAsnLeuGlyCysLysValValProAspCysAsnAsnTyr 67
|||||
165 CCAGAGTTGGGGAACATGGCTGCAAGGTGTTCCTGATTGTAACAACTA 214
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGly 84
|||||
215 CAGACAGAAGATCACCTCTGGATGGAGCGGATGTCAGATTCCCGGGG 264
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCCTGGACGGCGCAACTATATCTGGTGTGGTGCAGATGCCCT 314
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThr 117
|||||
315 AGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGTTAAACA 364
|||||
117 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGly 84
|||||
365 TATCAGAGCGCGCACTGAGAAAGGAGATTCAGGCCAGGAGTTAT 414
|||||
134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
415 CAGCCTACCAGGCTCCCTCCCGACCGCACACAGTGGCTTCCATCGCTAC 464
```

Thu Nov 15 10:52:10 2001

us-09-215-435-225.rnmp

```

151 GlnPhePheValThrLeuGlnGluGlyLysValIleSerLeuLeuProly 167
|||||
465 CAGTCTCTTCTATCTTCAGGAAGAAAGTCACTCTCTCTCTCCCA 514
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
|||||
515 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTGAACCGT 564
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCCACCTGGCGCAACCTGAAGCAGCACCAGTTCATGACCCAGAACTAC 614
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProly 217
|||||
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCGCAGCGAGCCAA 664
|||||
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
665 GCACAAAACACGCGGAGATAGTGCCTGC 695
|||||

seq_name: /cgn1_7/ptodata/1/pna/US6016_COMB.seq:US-60-169-629-7

seq_documentation_block:
; Sequence 7, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouquelieret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
US-60-169-629-7

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-60-169-629-7 ..

Align seg 1/1 to: US-60-169-629-7 from: 1 to: 826

1 MetGlyThrTrpMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTTGGCAATGAGCTGGTGCACAGCAGCAGCTGTTACTGGGTCTCAT 64
|||||
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGGTGGTCACTGGAGAGAGATGAGAACAGCCGCTGTGCCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCCTCCTGGACGAGGACACCTCTTTTGCAGGGCCCTTGAAGTTTCTAC 164
|||||

```

```

seq_name: /cgn1_7/ptodata/1/pna/US6018_COMB.seq:US-60-187-470-7

seq_documentation_block:
; Sequence 7, Application US/60187470
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouquelieret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: 78-US2.PRO
; CURRENT APPLICATION NUMBER: US/60/187,470
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
US-60-187-470-7

```

```
alignment_scores:
  Quality: 1224.00      Length: 227
  Ratio: 5.392          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-60-187-470-7 ..

Align seg 1/1 to: US-60-187-470-7 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTGGACAATGAGGTGGTCACAGCAGCACTGTTACTGGGTCTCAT 64
|||||
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGGTGGTCACTGGAGACGAGATGAGAACAGCCGTGTGCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCTCTCTGACGAGGACACCTCTTTTGGCAGGGCTTGAAGTTTCTAC 164
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
165 CCAGAGTTGGGAACATTGGCTGCAAGGTGTTCCTGATTGTAACAACATA 214
|||||
67 rArgGlnLysIleThrSerTrpMetClnProIleValLysPheProGlyA 84
|||||
215 CAGACAGAAGATCACCTCTGGATGGAGCGGATGAGTCAAGTTCCCGGGG 264
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCGTGACGGCGCAACCTATATCTCTGGTGTGATGGTGATCCAGATGCCCT 314
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
315 AGCAGACGACAAACCCACAGACAGATTTCTGGAGACATTTGGCTGGTAACAGA 364
|||||
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
365 TATCAAGGGCGCGACCTGAAGAAAGGAAGATTTCAGGGCCAGAGTTAT 414
|||||
134 exAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
415 CAGCCTACCAAGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 464
|||||
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProTy 167
|||||
465 CAGTTCTTTGTCTATCTTCAGGAAGGAAAGGTATCTCTCTCTCCCAA 514
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
515 GAAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTTCTGAACCGTT 564
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCACCTTGGCGGACCTGAAGCAGCAGCCAGCTTCATGACCCAGAACTAC 614
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProTy 217
|||||
615 CAGGACTCACCAACCTCCAGGCTCCACAGAAAGGCGCCAGCGAGGCCAA 664
|||||
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
665 GCACAAAACACAGGGGAGATAGTGCCTGC 695
```

```
seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-7

seq_documentation_block:
; Sequence 7, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
```

```
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Joubert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197.873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15...695
; NAME/KEY: sig_peptide
; LOCATION: 15...80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
US-60-197-873-7

alignment_scores:
  Quality: 1224.00      Length: 227
  Ratio: 5.392          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x US-60-197-873-7 ..

Align seg 1/1 to: US-60-197-873-7 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTGGACAATGAGGTGGTCACAGCAGCACTGTTACTGGGTCTCAT 64
|||||
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
65 GATGGTGGTCACTGGAGACGAGATGAGAACAGCCGTGTGCCATGAGG 114
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
115 CCTCTCTGACGAGGACACCTCTTTTGGCAGGGCTTGAAGTTTCTAC 164
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
165 CCAGAGTTGGGAACATTGGCTGCAAGGTGTTCCTGATTGTAACAACATA 214
|||||
67 rArgGlnLysIleThrSerTrpMetClnProIleValLysPheProGlyA 84
|||||
215 CAGACAGAAGATCACCTCTGGATGGAGCGGATGAGTCAAGTTCCCGGGG 264
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
265 CCGTGACGGCGCAACCTATATCTCTGGTGTGATGGTGATCCAGATGCCCT 314
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
315 AGCAGACGACAAACCCACAGACAGATTTCTGGAGACATTTGGCTGGTAACAGA 364
|||||
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
365 TATCAAGGGCGCGACCTGAAGAAAGGAAGATTTCAGGGCCAGAGTTAT 414
|||||
134 exAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
415 CAGCCTACCAAGCTCCCTCCACCGGCACACAGTGGCTTCCATCGCTAC 464
|||||
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProTy 167
|||||
465 CAGTTCTTTGTCTATCTTCAGGAAGGAAAGGTATCTCTCTCTCCCAA 514
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
515 GAAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTTCTGAACCGTT 564
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
565 TCACCTTGGCGGACCTGAAGCAGCAGCCAGCTTCATGACCCAGAACTAC 614
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProTy 217
|||||
615 CAGGACTCACCAACCTCCAGGCTCCACAGAAAGGCGCCAGCGAGGCCAA 664
|||||
217 sHisLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
665 GCACAAAACACAGGGGAGATAGTGCCTGC 695
```

Thu Nov 15 10:52:10 2001

us-09-215-435-225.rnmp

```

465 CAGTTCTTTGTCATCTTCAGGAAGGAAAGGTCACTCTCTCTCTCCGCCAA 514
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
515 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTTGAACCGTT 564
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
565 TCCACCTGGCGAACTGAAGCAAGCACCAGTTCATGACCCAGAACTAC 614
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLys 217
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCCAGGAGGCCAA 664
217 shisLysAsnGlnAlaGluLeuAlaLys 227
665 GCACAAAACACGCGGAGATAGTGCCTGC 695

```

seq_name: /cgnl_7/ptodata/1/pna/US6022_COMB.seq:US-60-224-006-7

```

seq_documentation_block:
; Sequence 7, Application US/60224006
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: ESTs encoding human proteins
; FILE REFERENCE: 95 US1.PRO
; CURRENT APPLICATION NUMBER: US/60/224,006
; CURRENT FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
; US-60-224-006-7

```

```

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:
US-09-215-435-225 x US-60-224-006-7 ..

```

Align seg 1/1 to: US-60-224-006-7 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeu 17
15 ATGGGTGGACAACTCGAGGCTCTTGGAAATGGACAGATTCTTGAACCGTT 64
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGlu 34
65 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCGCTGTGCCCATGAGG 114
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
115 CCCTCTGGACAGGACACCTCTTTGGCCAGGCTTGAAGTTTCTTAC 164
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnTyr 67
165 CCAGAGCTGGGAACATTGGCTGCAAGGTGTCTCTGATTGTAAACAATA 214
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84

```

```

215 CAGACAGAAGATCACCTCTCGGATGGAGCCGATAGTCAAGTCCCGGGG 264
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
265 CCGTGGAGCGCGCAACCTATATCTCGTGTATGGTGGATCCAGATGCCCT 314
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
315 AGCAGACAGAACCCACAGACAGAGATTCTGGAGACATTTGGCTGGTAACAGA 364
117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeus 134
365 TATCAAGGCGCCGACCTGAAGAAAGGAGGATTCAGGGCCAGGAGTTAT 414
134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
415 CAGCCTACGAGCTCCCTCCCGCCAGCCACACAGTGGCTTCCCATCGCTAC 464
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLys 167
465 CAGTTCTTTGTCATCTTCAGGAAGGAAAGGTCACTCTCTCTCTCCCAA 514
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
515 GGAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTTGAACCGTT 564
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
565 TCCACCTGGCGAACTGAAGCAAGCACCAGTTCATGACCCAGAACTAC 614
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLys 217
615 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCCAGGAGGCCAA 664
217 shisLysAsnGlnAlaGluLeuAlaLys 227
665 GCACAAAACACGCGGAGATAGTGCCTGC 695

```

seq_name: /cgnl_7/ptodata/1/pna/US6027_COMB.seq:US-60-278-258-12888

```

seq_documentation_block:
; Sequence 12888, Application US/60278258
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0010-1 P
; CURRENT APPLICATION NUMBER: US/60/278,258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 12888
; LENGTH: 904
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 204392.4
; US-60-278-258-12888

```

```

alignment_scores:
Quality: 1224.00 Length: 227
Ratio: 5.392 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225 x US-60-278-258-12888 ..

```

```
Align seg 1/1 to: US-60-278-258-12888 from: 1 to: 904
1 MetGlyTTPThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
89 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCAGCTGTACTGGGTCTCAT 138
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
139 GATGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 188
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuValPheTyr 50
189 CCCTCTTGGACAGACACCCCTCTTTGGCAGGGCCTTGAAGTTTCTAC 238
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
239 CCAGAGTTGGGGAACATGGCTGCAAGGTTGTCTCTGATGTGAACAATA 288
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
289 CAGACAGAAAGATCACCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 338
84 laValAspGlyAlaThrTyrlleLeuValMetValAspProAspAlaPro 100
339 CGGTGGACGGCGCAACCTATATCTCTGGTGGTGGATCCAGATGCCCT 388
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
389 ACAGACAGAACCCACAGACAGATTCGGAGACATTCGCTGGTGAACAGA 438
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeus 134
439 TATCAAGGGCGCCGACCTGAAGAAAGGAAGATTTCAGGGCCAGGAGTTAT 488
134 eAlaTyrlleGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
489 CAGCTTACCAAGCTCCCTCCCGCCAGCCGACACAGTGGCTTCCATCGCTAC 538
151 GlnPhePheValTyrlleGlnGluGlyLysValIleSerLeuLeuProLy 167
539 CAGTCTTTCTATCTTCAGGAAGAAAGTCATCTCTCTCTCTCTCCCAA 588
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
589 GGAACAAACAACTCGAGGCTCTTGAAATATGACAGATTCTTGAACCGTT 638
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
639 TCCACCTGGCGCAACCTGAAGCAAGCCACCCAGTTTCATGACCCAGAACTAC 688
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
689 CAGGACTCACCAACCTCCAGCTCCAGAGAAAGGGCCAGCGAGCCCAA 738
217 shiLysAsnGlnAlaGluIleAlaAlaCys 227
739 GCACAAACAAACAGCGGAGATAGTGCCTGC 769

seq_name: /cgnl_7/ptodata/1/pna/US097B_COMB.seq:US-09-726-176-1502

seq_documentation_block:
; Sequence 1502, Application US/09726176
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Donovan, Michael J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2018-001
; CURRENT APPLICATION NUMBER: US/09/726.176
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/167,841
; PRIOR FILING DATE: 1999-11-29
```

```
; NUMBER OF SEQ ID NOS: 1882
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1502
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-176-1502

alignment_scores:
Quality: 1220.00 Length: 227
Ratio: 5.374 Caps: 0
Percent Similarity: 100.000 Percent Identity: 99.559

alignment_block:
US-09-215-435-225 x US-09-726-176-1502
Align seg 1/1 to: US-09-726-176-1502 from: 1 to: 870
1 MetGlyTTPThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
71 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCAGCTGTACTGGGTCTCAT 120
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
121 GATGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 170
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuValPheTyr 50
171 CCCTCTTGGACAGGACACCCCTCTTTGGCAGGGCCTTGAAGTTTCTAC 220
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
221 CCAGAGTTGGGGAACATTCGCTGCAAGGTTGTCTCTGATGTGAACAATA 270
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
271 CAGACAGAAAGATCACCTCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 320
84 laValAspGlyAlaThrTyrlleLeuValMetValAspProAspAlaPro 100
321 CGGTGGACGGCGCAACCTATATCTCTGGTGGTGGATCCAGATGCCCT 370
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
371 ACAGACAGAACCCACAGACAGATTCGAGACATTCGCTGGTGAACAGA 420
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeus 134
421 TATCAAGGGCGCCGACCTGAAGGAAGGAAGATTTCAGGGCCAGGAGTTAT 470
134 eAlaTyrlleGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
471 CAGCTTACCAAGCTCCCTCCCGCCAGCCGACACAGTGGCTTCCATCGCTAC 520
151 GlnPhePheValTyrlleGlnGluGlyLysValIleSerLeuLeuProLy 167
521 CAGTCTTTCTATCTTCAGGAAGAAAGTCATCTCTCTCTCTCTCCCAA 570
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgp 184
571 GGAACAAACAACTCGAGGCTCTTGAAATATGACAGATTCTTGAACCGTT 620
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
621 TCCACCTGGCGCAACCTGAAGCAAGCCACCCAGTTTCATGACCCAGAACTAC 670
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
671 CAGGACTCACCAACCTCCAGCTCCAGAGAAAGGGCCAGCGAGCCCAA 720
217 shiLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
```

Thu Nov 15 10:52:10 2001

us-09-215-435-225.rnps

403 CCGTGGACGGCGCAACCTATATCTGTGTGGTGGTATCCAGATGCCCT 452
 101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
 453 AGCAGACAGAACCCAGACAGATTTCTGGAGACATTTGGTGTACAG 502
 117 PileLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
 503 TATCAAGGGCGCCACCTGAAGRAAGGAAGATTTCAGGGCCAGAGTTAT 552
 134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
 553 CAGCCTACAGGCTCCCTCCACCCGACACAGTGGCTTCCATCGCTAC 602
 151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProLy 167
 603 CAGTCTTTTGTCTATCTTCAGAAAGAAAGTCATCTCTCTCTTCCAA 652
 167 sGluAsnLysThrArgGlySerTrpLysMetaspArgPheLeuAsnArgP 184
 653 GGAACCAAACTCGAGGCTCTTGAAATGGACAGATTTCTGAACGTT 702
 184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
 703 TCCACCTGGCGGACCTGAAGCAAGCACCCAGTTCATGACCCAGACTAC 752
 201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
 753 CAGGACTCACCACCCCTCCAGGCTCCAGAGAAAGGGCCAGGAGCCAA 802
 217 shisLysAsnGlnAlaGluIleAlaAlaCys 227
 803 GCACAAAACCCAGCGGAGATAGCTGCCTGC 833

seq_name: /cgnl_7/ptodata/1/pna/US095D_COMB.seq:US-09-591-316-56

seq_documentation_block:
 ; Sequence 56, Application US/09591316
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 47 Human Secreted Proteins
 ; FILE REFERENCE: P2035P1
 ; CURRENT APPLICATION NUMBER: US/09/591.316
 ; CURRENT FILING DATE: 2000-06-09
 ; EARLIER APPLICATION NUMBER: PCT/US99/29950
 ; EARLIER FILING DATE: 1999-12-16
 ; EARLIER APPLICATION NUMBER: 60/113,006
 ; EARLIER FILING DATE: 1998-12-18
 ; EARLIER APPLICATION NUMBER: 60/112,809
 ; EARLIER FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 56
 ; LENGTH: 1028
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1022)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1026)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1027)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1028)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-591-316-56

721 GCACAAAACCCAGCGGAGATAGCTGCCTGC 751
 seq_name: /cgnl_7/ptodata/1/pna/PCTUS_COMB.seq:PCT-US99-29950-56
 seq_documentation_block:
 ; Sequence 56, Application PC/US9929950
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: 47 Human Secreted Proteins
 ; FILE REFERENCE: P2035.PCT
 ; CURRENT APPLICATION NUMBER: PCT/US99/29950
 ; CURRENT FILING DATE: 1999-12-16
 ; EARLIER APPLICATION NUMBER: 60/112,809
 ; EARLIER FILING DATE: 1998-12-17
 ; EARLIER APPLICATION NUMBER: 60/113,006
 ; EARLIER FILING DATE: 1998-12-18
 ; NUMBER OF SEQ ID NOS: 230
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 56
 ; LENGTH: 1028
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1022)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1026)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1027)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1028)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; PCT-US99-29950-56

alignment_scores:
 Quality: 1218.00 Length: 227
 Ratio: 5.389 Gaps: 0
 Percent Similarity: 99.559 Percent Identity: 99.559
 alignment_block:
 US-09-215-435-225 x PCT-US99-29950-56 ..
 Align seg 1/1 to: PCT-US99-29950-56 from: 1 to: 1028
 1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuLeuLeu 17
 153 ATGGGTGGCAATGAGCTGTGTACACAGCAGTGTCTGCTCAT 202
 17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGlu 34
 203 GATGGTGTCTACTGAGAGGAGGATGAGACAGCCGTGTGCCCATGAGG 252
 34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
 253 CCCTCTTGACAGGACACCCCTTTTGGCCAGGGCCTTGAAGTTTCTAC 302
 51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
 303 CCAGAGTGGGGAACATGGCTGCAAGTGTGCTGATTGTAAACAATA 352
 67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
 353 CACAGAGAAGATCACCTCTCTGATGAGCCGATAGTCAAGTCCCGGGG 402
 84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100


```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035p1
; CURRENT APPLICATION NUMBER: US/09/895,298
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29950
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/112,809
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1022)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1026)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1027)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1028)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-895-298-56

alignment_scores:
    Quality: 1218.00      Length: 227
    Ratio: 5.389          Gaps: 0
    Percent Similarity: 99.559      Percent Identity: 99.559

alignment_block:
US-09-215-435-225 x US-09-895-298-56

Align seg 1/1 to: US-09-895-298-56 from: 1 to: 1028

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
153 ATGGGTTGGACAATGAGCTGGTCACAGCAGCAGCTTTACTGGTCTCAT 202
|||||
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGlu 34
|||||
203 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGGCCCATGAGG 252
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
253 CCCTCTTGGACGAGCAGACCCCTCTTTGCCAGGCGCTTGAAGTTTCTAC 302
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
303 CCAGAGTTGGGAACATTTGGCTGCAAGGTGTCTCTGATTGTAACAATA 352
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
353 CAGACAGAGAAGATCACCTCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 402
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
403 CCGTGGACGGCCCACTATATCTCTGGTGGTGGATCCAGATGCCCTT 452
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
453 AGCAGAGCAGAACCCACAGACAGAGATTCGGAGACATTCGCTGGTAACAGA 502
|||||
117 pileLysGlyAlaAspLeuLysLysGlyLysIleGlnGlnGluLeuS 134
|||||
503 TATCAAGGGCGCCGACCTTGAAGAGGGAAGATTCAGGGCCAGGAGTTAT 552
|||||
134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
|||||
553 CAGCTACACAGGCTCCCTCCCGCCGACACAGAGTGGCTTCATCGCTAC 602
|||||
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
|||||
603 CAGTTCTTTGTCTATCTTCAGGAAGGAAAGTCATCTCTCTCTCCCA 652
|||||
167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
|||||
653 GGAAACAAACAACTCGAGGCTCTTGGAAATGGACAGATTCTCTGAACCGTT 702
|||||
184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
|||||
703 TCCACTTGGCCGAACCTGAAGCAAGCACCAGTTTATGACCCAGAACATAC 752
|||||
201 GlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluProLy 217
|||||
753 CAGGACTCACCAACCTCCAGGCTCCAGAGAAAGGCGCAGCGGCCCA 802
|||||
217 shiLysAsnGlnAlaGluIleAlaAlaCys 227
|||||
803 GCACAAAAACACGCGGAGATAGCTGCCTGC 833

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-895-298-56

seq_documentation_block:
; Sequence 56, Application US/09895298
; GENERAL INFORMATION:
```


THIS PAGE BLANK (USPTO)


```

471 ..GGGTGGCGCCAGAACTTCAACACTCGCGAGTTGCTGAGATCTACAAT 518

186 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
||||| ||| |||:||||: |||
519 CTCGGCTTCCGTCGCGCGAGTTTCTACAATTCTCAG 557

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-845-849-3

seq_documentation_block:
; Sequence 3, Application US/09845849
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: WEIGEL, Detlef
; APPLICANT: KARDAILSKY, Igor
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; FILE OF INVENTION: MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT
; FILE REFERENCE: SALKINS.026DV1
; CURRENT APPLICATION NUMBER: US/09/845,849
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/060,726
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-845-849-3

alignment_scores:
Quality: 164.00 Length: 113
Ratio: 2.412 Gaps: 4
Percent Similarity: 60.177 Percent Identity: 36.283

alignment_block:
US-09-215-435-225 x US-09-845-849-3/rev ..

Align seg 1/1 to reverse of: US-09-845-849-3 from: 1 to: 856

90 TyrIleLeuValMetValAspProGluAlaSerThrArgAlaGluProAr 106
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 TATACCTTTGGTTATGGTGATCCAGATGTTCCAGTGCTCAGCAACCTCA 553

106 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
552 CCTCCGAGAAATATCTCCATTGGTTGGTGACTGATATCCCTGCTACA... 507

123 euLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaPro 139
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
506 .....ACTGGAAACAACTTTGGCAATGAGATTGTGTACGAAATCCA 462

140 SerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 156
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
461 AGTCCCACTGCA.....GGAAATTCATCGTGTGTTATATTGTTTCG 418

156 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 173
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 ACAGCTTGGCAGCAACAGTGTATGCACCA..... 387

173 LysSerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 ..GGGTGGCGCCAGAACTTCAACACTCGCGAGTTTCTGAGATCTACAAT 339

186 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
398 CTCGSCCTTCCCGTGGCGCAGTTTCTACAATTCTCAG 300

seq_name: /cgnl_7/ptodata/1/pna/US60_NEW_COMB.seq:US-60-325-448-4131

seq_documentation_block:
; Sequence 4131, Application US/60325448
; GENERAL INFORMATION:

```

```

; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 4131
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-60-325-448-4131

alignment_scores:
  Quality: 158.00      Length: 141
  Ratio: 1.859         Gaps: 4
  Percent Similarity: 60.284  Percent Identity: 29.078

alignment_block:
US-09-215-435-225 x US-60-325-448-4131
Align seg 1/1 to: US-60-325-448-4131 from: 1 to: 891

76 GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 92
255 AAGCCACGCATTGAGTGGTGGGACATGAGATCCTTCTTCACTCT 304
92 uValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgP 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 GGTCAATGACGGACCCAGAGTGTCCAGGGGCTAGTATCCATACCTTAGGG 354
109 heTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 AGCATCTCCACTGGATCGTCAGTGATATCTCTGGCACC.....ACA 395
126 GlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaProSerProPr 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
396 GATGCTTCTTTTGAAGGAGGTGGTGGAGCTACGAGAGCCCAAGCCC... 443
142 alaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGlu 159
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
444 ...AACATGTGCATTACAGAGTTCCACCTTCGTGTGTTCACACAGAAGA 489
159 lYlValIleSerLeuLeuProLysGluAsnLysThrArgGlySerTrp 175
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
490 AGCGGAGGCCATGACCTCC.....TCCACAGGGACTACTTTC 530
176 LysMetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSe 192
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
531 AACACGCGCGCTTCGCCAACGAGAAGACCTTGGCCTCCCGGTGGCTGC 580
192 rThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrLeuGlnAlaP 209
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
581 CGTCTACTTCAACGGCGCAG.....CGGAGACAGCGGCACCGCCGCT 624
209 roArgGluAlaSerGluPro 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
625 GACGGAAACAATCAGCCCAACCC 647

seq_name: /cgnl_7/ptodata/1/pna/US60_NEW_COMB.seq:US-60-325-448-971

seq_documentation_block:
; Sequence 971, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 971

```

```

; LENGTH: 535
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-971

alignment_scores:
  Quality: 105.50      Length: 113
  Ratio: 1.598         Gaps: 2
  Percent Similarity: 58.407  Percent Identity: 26.549

alignment_block:
US-09-215-435-225 x US-60-325-448-971
Align seg 1/1 to: US-60-325-448-971 from: 1 to: 535

86 AspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerAr 102
|||:|||||:|||||:|||||:|||||:|||||:|||||:
185 GATCTTCTTCACTGGTGGGTTATGACAGACCCAGATGTGCCAGGACC 234
102 gAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleL 119
:::|||||:|||||:|||||:|||||:|||||:|||||:
235 AAGTGATCCTTATCTAAGGAGCACCCTTCATTCATTTACGCAGGATTGT 284
119 ysGlyAlaAspLeuLysLysGlyLysIleGlnGlnGluLeuSerAla 135
|||:|||||:|||||:|||||:|||||:|||||:|||||:
285 TACTGATATACCTGGGACAACAGGATGCTTTTTCGCGAGGTCATAAGC 334
136 TyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyrGlnPh 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
335 TATGAGAGTCCAAAGCG.....AACATTTGGCATCCATAGTTTCATTT 378
152 epheValTyrLeuGlnGluGlyLysValIleSerLeuProLysGluA 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 TGTGCTCTTCAACGACGAGCGGCAAACTGTAATTGTGCCA..... 421
169 snLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHis 185
:::|||||:|||||:|||||:|||||:|||||:|||||:
422 ...TCCTTACGGGACCATTTTCAACACCCCGGTTCCGCGAGGAGAAATGAT 469
186 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 CTGGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 508

seq_name: /cgnl_7/ptodata/1/pna/US60_NEW_COMB.seq:US-60-325-448-1413

seq_documentation_block:
; Sequence 1413, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 1413
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-1413

alignment_scores:
  Quality: 103.00      Length: 107
  Ratio: 1.537         Gaps: 4
  Percent Similarity: 62.617  Percent Identity: 30.841

alignment_block:
US-09-215-435-225 x US-60-325-448-1413
Align seg 1/1 to: US-60-325-448-1413 from: 1 to: 321

```


198 nAsnTyrGlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerG 215
 405AACCGCAGCGGGAGACAGCGGCAC 429

215 luProLysHisLysAsnGlnAla 222
 430 GCCGCGCTGACGGAACAATCA 452

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-6698

seq_documentation_block:

; Sequence 6698, Application US/09815242

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6698

; LENGTH: 951

; TYPE: DNA

; ORGANISM: Enterococcus faecalis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(951)

US-09-815-242-6698

alignment_scores:

Quality: 80.50 Length: 218

Ratio: 0.782 Gaps: 9

Percent Similarity: 47.248 Percent Identity: 20.642

alignment_block:

US-09-215-435-225 x US-09-815-242-6698 ..

Align seg 1/1 to: US-09-815-242-6698 from: 1 to: 951

33 GluAlaLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPh 49

||||| 358 GAAGCCTTACTAGCAGACACTGTGTAAATCGATGCTCGTAACGATTA 407

49 eTyrProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnA 66

||||| 408 TGAATATGATTAGTCAITTCGTCGTCGCCGCCAGATATCCGTA 457

66 snTyrArgGlnLysIleThrSerTrpMetGluProIleValLysPhePro 82

||||| 458 GCTTCGTCGAG...TTACCACATGATTCGC..... 486

83 GlyAlaValAspGlyAlaThrTyrIleLeuValMetValAspProAspAl 99
 486 486
 99 aProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValT 116
 487GAGACACAAGACAAATTTATGGATAAAAAAATTTGTTA 523
 116 hrAspIleLysGlyAlaAspLeuLysLysGlyIleGlnGlyGlnGlu 132
 524 CCTATTGTACTGGCGG...ATTCGCTGTGAAAAATTTCTGTGCTGTTA 570
 133 LeuSerAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisAr 149
 571 TTAAGAAGAGGATTTGAAGATGTTGCTCAATTCATGCTGTTATCGCCAA 620
 149 gTyr.....GlnPhePheValTyrLeuGlnGluGlyLysVal. 161
 621 CTATGGAAAAATCCAGAAACACGTCGCGCAACTTTGGACGCGCAAAATGT 670
 162IleSerLeuLeuProLysGluAsnLys 170
 671 ATGCTTTTGATGACCAATCAGTGTGCAATTAATCATGTTGATAAAAA 720
 171 ThrArgGly...SerTrp.....LysMetAspArgPheLeuAs 182
 721 GTTATGGGAAGACTGTTTGATGGACACCTTGGACGCTACATTA 770
 182 nArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnA 199
 771 C.....TGTCAAACCCAGATGTAATCGTCAAAATCTTAACITCAG 811
 199 snTyrGlnAspSerProThrLeu.....Gln 207
 812 AGAAAAATGAACATAAACATTTAGTGGCTGCTCATTAGATAGTAGCCAG 861
 208 AlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAlaGluIl 224
 862 CATCTGCCCAACCGTTATGTAATAATAATTAACAGAACAGCAGAGT 911
 224 eAla 225
 912 TGCT 915

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-6710

seq_documentation_block:

; Sequence 6710, Application US/09922340

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained

; FILE REFERENCE: 20411-726CON2

; CURRENT APPLICATION NUMBER: US/09/922,340

; CURRENT FILING DATE: 2001-08-03

; PRIOR APPLICATION NUMBER: 09/353,690

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: US 09/004,182

; NUMBER OF SEQ ID NOS: 12181

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6710

; LENGTH: 439

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-922-340-6710

alignment_scores:

Quality: 79.00 Length: 47

Ratio: 3.038 Gaps: 1

Percent Similarity: 55.319 Percent Identity: 36.170

```

alignment_block:
US-09-215-435-225 x US-09-922-340-6710  ..
Align seg 1/1 to: US-09-922-340-6710 from: 1 to: 439
112 HisTrpLeuValThrAspIleGlyAlaAspLeuLysGlyLysIle 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
12 CACTGGCTGCTAACCAACATCCGGGTAAACCGGGTG.....GC 49
128 eClnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProAlaHis 145
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 TGAAGGACAGGTGACGTGTCCTACCTCCGCCCTTCCTCCCGAGGCT 99
145 erGlyPheHisArgTyrGlnPheValTyrLeuGlnGlu 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 CCGGCATCCACCGTCTTCCCTCCTGCTCTTCAAGCAGGAC 140
seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-481

seq_documentation_block:
; Sequence 481, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 481
; LENGTH: 24961
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20774)..(21481)
US-09-897-516-481

alignment_scores:
Quality: 79.00 Length: 201
Ratio: 0.878 Gaps: 9
Percent Similarity: 44.776 Percent Identity: 20.398

alignment_block:
US-09-215-435-225 x US-09-897-516-481/rev ..
Align seg 1/1 to reverse of: US-09-897-516-481 from: 1 to: 24961
18 MetValThrGlyAspGluAspGluAsnSerProCysAlaHisGluAl 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22785 ATGATAATCACTTCACTCACTGAATTAAGAAATAT..... 22753
34 aLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyrP 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22752 .....AGTAAAGACGACGAGAAATTC 22725
51 roGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnTyr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22724 ATGAGTTGGTGGTAAACGGTGGTAAATCGTCACCTGAA..... 22687
68 ArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAl 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22686 .....CTTGCTGGACAGAT.....GGCC 22667

```

```
US-09-215-435-225 x US-09-897-516-3027/rev ..
Align seg 1/1 to reverse of: US-09-897-516-3027 from: 1 to: 24961

18 MetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluAl 34
||||:||||: ||| :|||:
22785 ATGATAATCACTTCAACTCAATTTAAGAATAAT..... 22753

34 aLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyrP 51
||||: ||| :|||
22752 .....GATTTCCTG.....AGTAAAGACGAGGAATTCA 22725

51 roGluLeuGlyAsnIleGlyCysLysValProAspCysAsnAsnTyr 67
||||:||||: ||| :|||
22724 ATGAGTTTGGTGAACGGTGGTAAATCGCTCACCTGAA..... 22687

68 ArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAl 84
::: :||| :|||
22686 .....CTTGCGCTGGACAGAT.....GCGCC 22667

84 aValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProS 101
||||: ||| :|||
22666 TCCCGGACAAAAGTTTTCGGGTACAGGTTTATGATCTGATGCCCAA 22617

101 erArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAsp 117
::: :||| :|||
22616 CGGGCAGT.....GGTTTCTG...CATTTGGTGCCATTTGAT 22582

118 Ile.....GATTTCCTG.....AGTAAAGACGAGGAATTCA 22725

22581 ATTCCGCTGGAATAACAGGCTTTATCCCGGATGCCGACGAGTATGG 22532

120 yAlaAspLeuLysLysGlyLysIleGlnGlyGlnGlu.....LeuS 134
||||: ||| :|||
22531 CAGAAATTCACGGTGGCATTATTCAGACAGAAATGATTACGGGCAGT 22482

134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
||||: ||| :|||
22481 TTGTTTCGGTGGACCTGCCCGCCAGAGGAGATAAACCCCATCGCTAT 22432

201 Gln 201
|||
22308 CAG 22306

seq_name: /cgnl_7/ptodata/1/pna/us09_NEW_COMB.seq:us-09-897-516-3028
```

```
seq_documentation_block:
; Sequence 3028, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897.516
; CURRENT FILING DATE: 2001-06-29
```

```
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 3028
; LENGTH: 24961
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2388)..(3563)
US-09-897-516-3028

alignment_scores:
Quality: 79.00 Length: 201
Ratio: 0.878 Gaps: 9
Percent Similarity: 44.776 Percent Identity: 20.398

alignment_block:
US-09-215-435-225 x US-09-897-516-3028/rev ..
Align seg 1/1 to reverse of: US-09-897-516-3028 from: 1 to: 24961

18 MetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluAl 34
||||:||||: ||| :|||:
22785 ATGATAATCACTTCAACTCAATTTAAGAATAAT..... 22753

34 aLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyrP 51
||||: ||| :|||
22752 .....GATTTCCTG.....AGTAAAGACGAGGAATTCA 22725

51 roGluLeuGlyAsnIleGlyCysLysValProAspCysAsnAsnTyr 67
||||:||||: ||| :|||
22724 ATGAGTTTGGTGAACGGTGGTAAATCGCTCACCTGAA..... 22687

68 ArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAl 84
::: :||| :|||
22686 .....CTTGCGCTGGACAGAT.....GCGCC 22667

84 aValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProS 101
||||: ||| :|||
22666 TCCCGGACAAAAGTTTTCGGGTACAGGTTTATGATCTGATGCCCAA 22617

101 erArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAsp 117
::: :||| :|||
22616 CGGGCAGT.....GGTTTCTG...CATTTGGTGCCATTTGAT 22582

118 Ile.....GATTTCCTG.....AGTAAAGACGAGGAATTCA 22725

22581 ATTCCGCTGGAATAACAGGCTTTATCCCGGATGCCGACGAGTATGG 22532

120 yAlaAspLeuLysLysGlyLysIleGlnGlyGlnGlu.....LeuS 134
||||: ||| :|||
22531 CAGAAATTCACGGTGGCATTATTCAGACAGAAATGATTACGGGCAGT 22482

134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
||||: ||| :|||
22481 TTGTTTCGGTGGACCTGCCCGCCAGAGGAGATAAACCCCATCGCTAT 22432

151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProly 167
||||: ||| :|||
22431 ATTTTCACGGTACATGCCATGCGGTAGAAAAAACTGGGTATT..... 22390

167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
::: :||| :|||
22389 .....AACGCTGAAATAACAAATGCGTCACACGTT 22359

184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
||| :||| :|||
22358 TTATGATCTACGCCAATAGTCTGGCAACGCCCACTATTACAGGCTATTAC 22309

201 Gln 201
|||
```



```
68 ArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAl 84
   :::||| :::
22686 .....CTTGCCTGGACAGAT.....GCGCC 22667

84 aValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProS 101
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
22666 TCGCGGAACAAAAGTTTTCGGGTACGGTTTATGATCCTGATGCCCCAA 22617

101 erArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAsp 117
   : : : : : : : : : : : : : : : : : : : : : : : :
22616 CGGGCAGT.....GGTTTCTGG...CATTGGGTGGCATTGTAT 22582

118 Ile.....LysG1 120
   |||
22581 ATTCCGCTGGAATACAGGCTTTATCCCGGATGCCGGACGGAGTGATGG 22532

120 yAlaAspLeuLysLysGlyLysIleGlnGlnGlu.....LeuS 134
   :::: ||| ||| ||| :::: :::: :::: :::: :::: ::::
22531 CAGCAAAATTACCGGCTGGCATCATTCAGAGCAGAAATGATTACGGGCAGT 22482

134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
   :::: :::: :::: :::: :::: :::: :::: :::: ::::
22481 TTGGTTTCGGTGGAGCTGCCGCCAGAGGAGATAAGCCCATCGCTAT 22432

151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLy 167
   ||| ||| :::: :::: :::: :::: :::: :::: ::::
22431 ATTTTCAGGTACATGCCATGTCGGTAGAAAAACTGGGTATT..... 22390

167 sGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgP 184
   : : : : : : : : : : : : : : : : : : : : : :
22389 .....AACGCTGAAATAACAAATGCGGTCACACGTT 22359

184 heHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyr 200
   || :::: :::: :::: :::: :::: :::: ::::
22358 TTATGATCTACGCCCAATAGTCTGGCAACCCCACTATTACAGGCTATTAC 22309

201 Gln 201
   |||
22308 CAG 22306
```

THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-225 to: EST:* out_format : pfs
Date: Nov 14, 2001 10:31 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnl_1/USPto_spool/US09215435/runat_14112001_141104_2118/app_query.fasta_1.1519
-DB=EST -QPM=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09215435 -CGN1_1_6733 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225
Query length: 227
Database: EST:*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 5896.450000

score_list:

Sequence	Strd	Orig	Zscore	Escore	Len	Documentation
gb_est102:BG545668	+	1028	0.90	2.9e-107	749	1
gb_est103:BF854481	+	921	0.00	3.8e-95	564	1
gb_est103:BF852989	+	884	0.00	7.1e-91	603	1
gb_est103:BF852995	+	836	0.00	2.2e-87	563	1
gb_est103:BF853000	+	853	0.00	1.9e-85	558	1
gb_est103:BF853920	+	792	0.00	1.8e-80	539	1
gb_est103:BF853928	+	785	0.00	1.2e-79	536	1
gb_est103:BF846063	+	763	0.00	3.3e-77	464	1
gb_est103:BF848191	+	735	0.00	4.5e-74	463	1
gb_est103:BF846065	+	725	0.00	6.4e-73	467	1
gb_est103:BF848190	+	723	0.00	1.5e-72	463	1
gb_est103:BF846175	+	718	0.00	1.8e-72	547	1
gb_est103:BF857902	+	710	0.00	5.8e-72	547	1
gb_est103:BF844451	+	708	0.00	1.4e-71	523	1
gb_est103:BF846059	+	705	0.00	1.4e-70	462	1
gb_est103:BF846059	+	699	0.00	8.7e-70	552	1
gb_est103:BF846059	+	695	0.00	2.4e-69	534	1
gb_est103:BF846558	+	687	0.00	1.7e-68	532	1
gb_est103:BF846558	+	686	0.00	2.0e-68	451	1
gb_est103:BF846558	+	675	0.00	4.3e-67	511	1
gb_est103:BF846558	+	672	0.00	7.2e-67	463	1
gb_est103:BF846558	+	627	0.00	1.0e-61	430	1
gb_est103:BF846558	+	626	0.00	1.4e-61	452	1
gb_est103:BF846558	+	611	0.00	1.4e-59	789	1
gb_est103:BF846558	+	605	0.00	3.0e-59	395	1
gb_est103:BF846558	+	604	0.00	4.3e-59	423	1
gb_est103:BF846558	+	582	0.00	1.6e-56	459	1
gb_est103:BF846558	+	581	0.00	1.3e-56	334	1
gb_est103:BF846558	+	560	0.00	4.3e-54	407	1
gb_est103:BF846558	+	556	0.00	1.5e-53	457	1
gb_est103:BF846558	+	555	0.00	6.2e-53	1048	1
gb_est103:BF846558	+	540	0.00	7.5e-52	413	1
gb_est103:BF846558	+	529	0.00	1.5e-50	402	1
gb_est103:BF846558	+	521	0.00	1.5e-49	454	1
gb_est103:BF846558	+	491	0.00	3.3e-46	480	1
gb_est103:BF846558	+	490	0.00	2.5e-46	280	1
gb_est103:BF846558	+	487	0.00	9.7e-46	412	1
gb_est103:BF846558	+	482	0.00	7.3e-45	440	1
gb_est103:BF846558	+	481	0.00	2.7e-45	277	1
gb_est103:BF846558	+	479	0.00	8.7e-45	436	1
gb_est103:BF846558	+	474	0.00	2.9e-44	400	1

gb_est103:BE165776 + 449.00 907.18 2.1e-41 401 ! BE165776 CM3-HT0486-310100-0
gb_est105:H83927 - 444.00 898.38 6.4e-41 352 ! H83927 ys64b02.sl Soares ret
gb_est110:W33189 - 444.00 895.97 8.7e-41 437 ! W33189 zc09f01.sl Soares par
gb_est104:F31628 + 436.00 885.42 3.4e-40 258 ! F31628 HSPD22948 HM3 Homo sa
seq_name: gb_est102:BG545668
seq_documentation_block:
LOCUS BG545668 749 bp mRNA EST 04-APR-2001
DEFINITION 602572933F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4700840 5',
mRNA sequence.
ACCESSION BG545668
VERSION BG545668.1 GI:13544333
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 749)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI535 row: b column: 09
High quality sequence stop: 740.
FEATURES
Source
Location/Qualifiers
1..749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4700840"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNM-LiB (Clontech); Site:1:
SfiI (ggccattggcc); Site:2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGAGCGGCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 184 a 206 c 205 g 154 t
ORIGIN
alignment_scores:
Quality: 1028.50 Length: 224
Ratio: 4.851 Gaps: 5
Percent Similarity: 94.643 Percent Identity: 93.304
alignment_block:
US-09-215-435-225 x BG545668 ..
Align seg 1/1 to: BG545668 from: 1 to: 749
1 MetGlyTrpThrMetArgLeuValThrAlaLaLeuLeuGlyLeuLeu 17
|||||
86 ATGGGTGGACAATGAGGTGGTGTACAGCAGCAGCTGTACTGGTCTCAT 135
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGlu 34
|||||
136 GATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuValPheTyr 50

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-271200-005-d12&t3=2000-12-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 562.

FEATURES

Location/Qualifiers

1..564

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="EN0093"

/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 149 a 156 c 141 g 118 t

ORIGIN

alignment_scores:
Quality: 921.00 Length: 183
Ratio: 5.174 Gaps: 1
Percent Similarity: 97.268 Percent Identity: 94.536

alignment_block:
US-09-215-435-225 x BF854481 ..

Align seg 1/1 to: BF854481 from: 1 to: 564

17 MetMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisG1 33
16 CTGTTTGTGACCCATCATGAGACCCAGTAACAGCCCGTGTGCCATGA 65
33 uAlaLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheT 50
66 GGCCCTCTTGGACGAGCACCCCTCTTTTCCAGGGCCCTTGAAGTTTCT 115
50 yrProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsn 66
116 ACCCAGAGTTGGGAAACATTGGCTGCAAGGTTGTTCTGATTGTAACAAC 165
67 TyrArgGlnLysIleThrSerTrpMetGluProIleValLysPheProG1 83
166 TACAGACAGAAGATCACCTCTCTGGATGGAGCGGATAGTCAAGTTCCCGG 215
83 yAlaValAspGlyAlaThrTyrIleLeuValMetValAspProAspAla. 99
216 GGCCGTGGACGGCGCAACCTATATCTGTTGTTGGTGGATCCAGATGCC 265
100 ProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValTh 116
266 CCTAGCAGACAGCAACCCAGACAGACAGATTCTGGAGACATTGGCTGCTAAC 315
116 rAspIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluL 133
316 AGATATCAAGGGCGCCGACCTGAAGAAAGGAAGATTACAGGGCCAGGAGT 365
133 euSerAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArg 149
366 TATCAGCCTACAGGCTCCCTCCCGCCGACACAGTGGCTTCCATCGC 415
150 TyrGlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuPr 166
416 TACCAGTTCTTGTCTATCTTCAGGAAGAAAGATCATCTCTCTCTCTCC 465

186 CCCTCTTGACGAGGACACCTCTTTTGCCA.GGCCTTGAAGTTTCTTAC 234
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
235 CCAGAGTTGGGAAACATTGGCTGCAAGGTTGTCTCTGATTGTAACAATA 284
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
285 CACACAGAGATCACCTCTGGATGGAGCGGATAGTCAAGTTCCCGGGG 334
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
335 CCGTGGACGGCGCAACCTATATCTCTGTTGTTGGATCCAGATGCCCT 384
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
385 AGCAGACAGCAACCCAGACAGATTCTGGAGACATTGGCTGGTAACACA 434
117 pIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuS 134
435 TATCAAGGGCGCCGACCTGAAGGAAGGAAGATTACAGGGCCAGGATTAT 484
134 erAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
485 CAGCCTACAGGCTCCCTCCCGCCGACACAGTGGCTTCCATCGCTAC 534
151 GlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeuProL 167
535 CAGTTCTTTGTCATCTTCAGGAAGGAACAGTATCTCTCTCTCTCCCA 584
167 ysGluAsnLysThrArgGlySerTrpLysMetAspArg....PheLeuAsn 182
585 AGGAAAAACAACCTCGAGGCTCTTGGAAACATAGCACAGATTCTTGAAC 634
183 ArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMet.ThrGlnA 199
635 CGGTTCCACCTGGCGAACCTGTAGCAAGCAACCCAGTTCTATGAACCCAGA 684
199 sn.TyrGlnAspSer.ProThrLeuGlnAlaProArgGluArgAlaSerG 215
685 ACTTACAGGAGTCAACCAACCTCCAGGCTCCAGAGGAAGGAGCCAGCG 734
215 luProLysHisLys 219
735 AGCCAGACACCAAAA 748

seq_name: gb_est93:BF854481

seq_documentation_block:
LOCUS BF854481 564 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-271200-005-d12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF854481
VERSION BF854481.1 GI:12242225
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

JOURNAL MEDLINE COMMENT


```

166 olysGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsnA 183
|||||
466 CAAGGAAACAAACATCGAGGCTTTGGAAATGGACAGATTCTGAACC 515
|||||
183 rghPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGln 198
|||||
516 GTTTCACCTGGCGGAAGTGAAGCAAGCACCAGTTTCATGACCCAG 562

seq_name: gb_est93:BF852989

seq_documentation_block:
LOCUS BF852989 603 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-a12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852989
VERSION BF852989.1 GI:12240837
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 603)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-a12&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 560.
FEATURES
source
location
1..603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 140 a 157 c 171 g 135 t
ORIGIN

alignment_scores:
Quality: 884.00 Length: 171
Ratio: 5.293 Gaps: 0
Percent Similarity: 97.661 Percent Identity: 97.076

alignment_block:
US-09-215-435-225 x BF852989 ..
Align seg 1/1 to: BF852989 from: 1 to: 603

```

```

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuLeuLeuMe 17
|||||
91 ATGGGTGGACAATGAGCTGGTGCACAGCAGCAGCTTTACTGGGTCTCAT 140
|||||
17 MetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
141 GATGGTGGTCACTGGAGACGAGGATGAGAACGCCCTGTGCCCATGAGG 190
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
191 CCTCTTGAACGAGGACACCTCTTTTGCAGGGCTTGAAGTTTCTAC 240
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
241 CCAGAGTTGGGGAACATTGGCTGCAAGGTGTTCCTGATTGTAACAAC 290
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
291 CAGACAGAAGATCACCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGG 340
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||
341 CGGTGGACGGCGCAACCTATATCTCTGTGATGGTGGATCCAGATGCCCT 390
|||||
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
391 AGCAGAGCAGAGACCCAGACAGAGATTCTGGAGACATTGGCTGGTAACA 440
|||||
117 pIleLysGlyAlaAspLeuLysLysLysLysLysLysLysLysLys 134
|||||
441 TATCAAGGGCGCCGACCTCTGAAGAAAGGAAGATTTCAGGGCGAGGATT 490
|||||
134 exAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyr 150
|||||
491 CAGCTACACGAGCTCCTCCACCGGCACACAGTGGCTTCCATCCCTAA 540
|||||
151 GlnPhePheValTyrLeuGlnGlyLysValIleSerLeuLeuProLy 167
|||||
541 CAGTTCTTTGTCTATCTTCAGGGAGGAAAGTTCATCTCTCTCTTTCCAA 590
|||||
167 sGluAsnLysThr 171
|
591 GGGAAACAAAACT 603

seq_name: gb_est93:BF852995

seq_documentation_block:
LOCUS BF852995 563 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-e01 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852995
VERSION BF852995.1 GI:12240843
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 563)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

```

```

541 CAGTCTTTGCTATCTTCAG 561
seq_name: gb_est93:BF853000

seq_documentation_block: 558 bp mRNA EST 16-JAN-2001
LOCUS BF853000
DEFINITION MR2-EN0093-191200-001-f12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853000
VERSION BF853000.1 GI:12240848
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 558)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., de Oliveira, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
{http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-
191200-001-f12&t3=2000-12-19&t4=1}
Seq primer: puc 18 forward
High quality sequence stop: 558.

FEATURES
source
1..558
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 126 a 148 c 164 g 125 t
ORIGIN

alignment_scores:
Quality: 853.00 Length: 157
Ratio: 5.433 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225 x BF852995 ..
Align seg 1/1 to: BF852995 from: 1 to: 563

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
91 ATGGTTGGACAATGAGCTGGTCACAGCAGCACTGTACTGGGTCTCAT 140
17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
141 GATGGTGTCTACTGAGACGAGGATGAGACAGCCGCTGTGCCCATGAGG 190
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
191 CCCCTTGGACGAGACACCTCTTTTCCAGGCCCTTGAGTTTCTAC 240
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
241 CCAGAGTTGGGAACATTTGGCTGCAAGGTTTCTCTGATTGTACAACATA 290
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
291 CAGACAGAGATCATCCTCTGGATGGAGCCGATAGTCAAGTCCCGGGG 340
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
341 CCGTGGACGGCCACCATATATCTGGTGATGGTGGATCCAGATGCCCT 390
101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
391 AGCAGAGCAGAACCCACACAGAGATTCTGGAGACATTGGCTGGTAACAGA 440
117 pIleLysGlyAlaAspLeuLysLysLysIleGlnGlyGlnGluLeus 134
441 TATCAAGGCGCGGACCTATATCTGGTGATGGTGGATCCAGATGCCCT 490
134 exAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
491 CAGCCTACAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTAC 540
151 GlnPheValTyrLeuGln 157
|||||

```

```
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
193 CCTCTTGGACGAGACACCCCTTTTGGCAGGCGCTTGAAGTTTCTTAC 242

51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
243 CCAGAGTTGGGAACATTTGGCTGCAAGGTTGTCTCTGATTGTAACAATA 292

67 rArgGlnLysIleThrSerTyrMetGluProIleValLysPheProGlyA 84
|||||
293 CAGACAGAGATCACCTCTCTGATGGAGCGGATGATCAAGTTCCCGGGG 342

84 laValaspGlyAlaThrTyrIleLeuValMetValaspProaspAlaPro 100
|||||
343 CCGTGGACGGCGCAACCTATATCTCTGATGGAGCGGATGATCAAGTTCCCGGGG 342

101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
393 AGCAGACAGAACCCAGACAGAGATTTGAGACATTTGGCTGTTACACA 442

117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
443 TATCAAGGCGCGGACCTCTGAGAAAGGAGGATTCAGGNCAGGAGTTAT 492

134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyr 150
|||||
493 CAGCCTACCAAGGCTCCCTCCCGAGCGGCACACAGTGGCTTCATCGCTAC 542

151 GlnPhePheValTyr 155
|||||
543 CAGTCTTTGTCTAT 557

seq_name: gb_est93:BF853920

seq_documentation_block: 499 bp mRNA EST 16-JAN-2001
LOCUS BF853920
DEFINITION MR2-EN0093-261200-004-d01 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853920
VERSION BF853920.1 GI:12241664
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2at2-MR2-EN0093-
261200-004-d01&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 227.
Location/Qualifiers
1. .499

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site:1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 119 a 129 c 148 g 103 t
ORIGIN
alignment_scores:
Quality: 792.00 Length: 149
Ratio: 5.351 Gaps: 0
Percent Similarity: 99.329 Percent Identity: 97.987
alignment_block:
US-09-215-435-225 x BF853920
Align seg 1/1 to: BF853920 from: 1 to: 499
1 MetGlyTyrThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
48 ATGGGTGGACAAATGAGGCTGGTCACAGCAGCACTGTACTGGTCTCAT 97

17 tMetValValThrGlyAspGluAspGluAsnSerProCysAlaHisGluA 34
|||||
98 GATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGAGG 147

34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
148 CCTCTTGGACGAGGACACCCCTTTTGGCAGGCGCTTGAAGTTTCTTAC 197

51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
198 CCAGAGTTGGGAACATTTGGCTGCAAGGTTGTCTCTGATTGTAACAATA 247

67 rArgGlnLysIleThrSerTyrMetGluProIleValLysPheProGlyA 84
|||||
248 CAGACAGAGATCACCTCTCTGATGGAGCGGATGATCAAGTTCCCGGGG 297

84 laValaspGlyAlaThrTyrIleLeuValMetValaspProaspAlaPro 100
|||||
298 CCGTGGACGGCGCAACCTATATCTCTGATGGAGCGGATGATCAAGTTCCCGGGG 347

101 SerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
348 AGCAGACAGAACCCAGACAGAGATTTGAGACATTTGGCTGTTACACA 397

117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuS 134
|||||
398 TATCAAGGCGCGGACCTCTGAGAAAGGAGGATTCAGGNCAGGAGTTAT 447

134 erAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisArg 149
|||||
448 CAGGCTTACCAAGGCTCCCTCCCGAGCGGCACACAGTGGCTTCCATCGC 494

seq_name: gb_est93:BF853928

seq_documentation_block: 536 bp mRNA EST 16-JAN-2001
LOCUS BF853928
DEFINITION MR2-EN0093-261200-004-g12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853928
VERSION BF853928.1 GI:12241672
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE
AUTHORS
1 (bases 1 to 536)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brantani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2at2-MR2-EN0093-
261200-004-g12t3-2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 497.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT .121 a 139 c 164 g 112 t
ORIGIN
alignment_scores:
Quality: 785.00 Length: 148
Ratio: 5.340 Gaps: 0
Percent Similarity: 99.324 Percent Identity: 97.973
alignment_block:
US-09-215-435-225 x BF853928 ..
Align seg 1/1 to: BF853928 from: 1 to: 536
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
91 ATGGGTGGACCAATGAGGCTGCTCAGCAGCAGCTGTTACTGGGCTCAT 140
|||||
17 MetValValThrGlyAspGluAsnSerProCysAlaHisGluA 34
|||||
141 GATGGTGGTCACTGAGCAGCAGATGAGACAGCCGCTGTCCTATGAGG 190
|||||
34 laLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPheTyr 50
|||||
191 CCCTCTGGACGAGGACCCCTCTTTTGGCAGGGCCTTGAAGTTTCTAC 240
|||||
51 ProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTy 67
|||||
241 CCAGAGTTGGGGAACATTCGCTGCAAGGTCGTTCTGATTTGTAACAATA 290
|||||
67 rArgGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyA 84
|||||
291 CACAGAGAAGATCACCTCCTCGGATGGAGCGCATAGTCAAGTCCCGGGG 340
|||||
84 laValAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaPro 100
|||||

|||||
341 CCGTGGAGCGCCGACCTATATCTCTGGTATGCTGGATCCAGATGCCCT 390
|||||
101 SetArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAs 117
|||||
391 AGCAGACAGAACCCACAGACAGATTCTGGAGACATTTGGCTGGTAACAGA 440
|||||
117 pIleLysGlyAlaAspLeuLysLysGlyLysIleGlnGlnGlnLeuS 134
|||||
441 TATCAAGGGCGCCGACCTGAAGAAAGGGAAGATTCAGGGGCGAGGTTAT 490
|||||
134 erAlaIaTyrGlnAlaProSerProProAlaHisSerGlyPheHis 148
|||||
491 CAGGCTAACAGGGTCCCTCCGCCGACACACAGCTGCTTCCAT 534
|||||
seq_name: gb_est93:BF846063
seq_documentation_block:
LOCUS BF846063 464 bp mRNA EST 16-JAN-2001
DEFINITION QVO-EN0055-181000-439-f11 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF846063
VERSION BF846063.1 GI:12233213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brantani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QVO&t2=QVO-EN0055-
181000-439-f11&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 464.
Location/Qualifiers
1..464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"
/note="Organ: lung_normal; vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 134 c 122 g 114 t
ORIGIN
alignment_scores:
Quality: 763.00 Length: 142
Ratio: 5.373 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225 x BF846063/rev ..

Align seg 1/1 to reverse of: BF846063 from: 1 to: 464

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
444 ATGAGGCTGGTCACAGCAGCACTGTTACTGTGCTCATGATGGTGCAC 395
21 rGlyAspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspG 38
|||||
394 TGGAGACGAGGATGAGAACAGCCGCTGTGCCATGAGGCCCTCTTGGAGC 345
38 luAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGly 54
|||||
344 AGGACACCCCTCTTTTCCAGGCCCTTGAAGTCTTCTACCCAGATGGGG 295
55 AsnileGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysI 71
|||||
294 AACATTGGCTGCAAGGTTCTCTGATTGTAACAACACTACAGACAGAAGAT 245
71 eHrSerTrpMetGluProIleValLysPheProGlyAlaValAspGlyA 88
|||||
244 CACCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGGTGGACGGC 195
88 laThrTyrIleLeuValMetValAspProAspAlaProSerArgAlaGlu 104
|||||
194 CACCTTATATCTCTGGTATGGTGGATCCAGATGCCCTAGCAGACGAGAA 145
105 ProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAl 121
|||||
144 CCAGACAGAGATCTCGGAGACATTTGGCTGGTAAACAGATATCAAGGGCC 95
121 aAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnA 138
|||||
94 CGACCTGAAGAAAGGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACCA 45
138 laProSerProProAlaHisSerGly 146
|||||
44 CTCCTCTCCCGCCGACACAGTGGC 19

seq_name: gb_est93:BF848191

seq_documentation_block: 463 bp mRNA EST 16-JAN-2001
LOCUS BF848191 QV0-EN0055-031100-480-cl0 EN0055 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF848191

VERSION BF848191.1 GI:12235341

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 463)

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

seq_name: gb_est93:BF846065

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0st2-QV0-EN0055-
031100-480-cl0st3-2000-11-03st4-1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 463.

FEATURES

source

1..463
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 119 c 136 g 92 t
ORIGIN

alignment_scores:

Quality: 735,50 Length: 145
Ratio: 5,180 Gaps: 1
Percent Similarity: 97,931 Percent Identity: 95,172

alignment_block:

US-09-215-435-225 x BF848191 ..

Align seg 1/1 to: BF848191 from: 1 to: 463

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
30 ATGAGGCTGGTCACAGCAGCA...GTACTGGGTCTCATGATGGTGTCTAC 76
21 rGlyAspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspG 38
|||||
77 TGGAGACGAGGATGAGAACAGCCCGTGTGCCATGAGGCCCTCTTGGAGC 126
38 luAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGly 54
|||||
127 AGGACACCCCTCTTTCCAGGGCCCATGAAGTTTCTACCCAGAGTTGGGG 176
55 AsnileGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysI 71
|||||
177 AACATTGGCTGCAAGGTTGTCTCTGATTGTAACTACAGACAGAAGAT 226
71 eHrSerTrpMetGluProIleValLysPheProGlyAlaValAspGlyA 88
|||||
227 CACCTCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGGTGGACGGC 276
88 laThrTyrIleLeuValMetValAspProAspAlaProSerArgAlaGlu 104
|||||
277 CACCTTATATCTCTGGTATGGTGGATCCAGATGCCCTAGCAGACGAGAA 326
105 ProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAl 121
|||||
327 CCAGACAGAGATCTCGGAGACATTTGGCTGGTAAACAGATATCAAGGGCG 376
121 aAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnA 138
|||||
377 CGACCTGAAGAAAGGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACCA 426
138 laProSerProProAlaHisSerGlyPheHisArg 149
|||||
427 CTCCTCTCCCGCCGACACACAGTGGCATCGATAAA 461

```

seq_documentation_block: 457 bp mRNA EST 16-JAN-2001
LOCUS BF846065
DEFINITION QV0-EN0055-181000-439-g06 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF846065
VERSION BF846065.1 GI:12233215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV0st2-QV0-EN0055-
181000-439-g06&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 467.
FEATURES
source
1. 467
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 134 c 126 g 113 t
ORIGIN

alignment_scores:
Quality: 725.50 Length: 143
Ratio: 5.182 Gaps: 2
Percent Similarity: 97.902 Percent Identity: 97.203

alignment_block:
US-09-215-435-225 x BF846065/rev ..
Align seg 1/1 to reverse of: BF846065 from: 1 to: 467
5 MetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetValValrh 21
|||||
447 ATGAGGTCGTGCACAGCAGCAGCTGTACTGGGTCATCATGATGGTGCAC 398
|||||
21 rGlyAspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspG 38
|||||
397 TGGAGACGAGGATGAGAACAGCCGTGTGCCCATGAGGCCCTCTTGGAGC 348
|||||
38 LuAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGly 54
|||||

```

```

347 AGGACACCCCTCTTTGCCAGGGCCCTTCAAGTCTTCTACCCAGAGTTGGG 298
55 AsnIleGlyCysIysValValProAspCysAsnAsnTyrArgGlnIysI 71
|||||
297 AACATCGCTGCAAGGTGTCTCTGATTTGAACAACATACAGACAGAGAT 248
|||||
71 eThrSerTrpMetGluProIleValIysPheProGlyAlaValAspGlyA 88
|||||
247 CACCTCTGTGATGAGCGGATAGTCAAGTTCCTCCGGGGCCCTGACGGCG 198
|||||
88 laThrTrpIleLeuValMetValAspProAspAlaProSerArgAlaGlu 104
|||||
197 CAACCTATATCTGCTGATGCTGATCCAGATGCCCTAGCAGAGCAGAA 148
|||||
105 ProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleIysGlyAl 121
|||||
147 CCCAGACAGAGATTCTCGAGACATTGGCTGTAACAGATATCAAGGGCGC 98
|||||
121 aAspLeuIysIysGlyIysIleGlnGlyGlnGluLeuSerAlaTyrGlnA 138
|||||
97 GCACCTGAAGAAGGAGAGATTTCAGGGCCAGGAGTATATCAGCCTTAC...C 51
|||||
138 la.ProSerProAlaHisSerGly 146
|||||
50 GCTCCTCTCCCAACCGGCACACGGTGGC 24
|||||
seq_name: gb_est93:BF848190
seq_documentation_block: 463 bp mRNA EST 16-JAN-2001
LOCUS BF848190
DEFINITION QV0-EN0055-031100-480-c05 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF848190
VERSION BF848190.1 GI:12235340
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 463)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV0st2-QV0-EN0055-
031100-480-c05&t3=2000-11-03&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 463.
FEATURES
source
1. 463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application

```

No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 117 a 120 c 136 g 90 t

alignment_scores:
Quality: 723.50 Length: 146
Ratio: 5.095 Gaps: 2
Percent Similarity: 97.260 Percent Identity: 94.521

alignment_block:

US-09-215-435-225 x BF848190 ..

Align seg 1/1 to: BF848190 from: 1 to: 463

5 MetArgLeuValThrAlaAla..LeuLeuLeuGlyLeuMetMetValValt 21
|||||
29 ATGAGGCTGGTCACAGCAGCGTA...CTGGGTCTCATGATGGTGTCA 75
21 hrClyAspLysValSerProCysAlaHisGluAlaLeuLeuAsp 37
|||||
76 CTGGAGACGAGGATGAGAACGCCGTGTGCCCATGAGGCCCTCTGGAC 125
38 GluAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuG 54
|||||
126 GAGGACACCTCTTTTGCAGGACCTTGAAGTCTACTACCCAGAGTTGG 175
54 yAsnIleGlyCysLysValValProAspCysAsnTyrArgGlnLysI 71
|||||
176 GAACATTGGCTCAAGGTGTCTCTGATTGTAACAACATACACAGAGA 225
71 leThrSerTrpMetGluProIleValLysPheProGlyAlaValaspGly 87
|||||
226 TCACCTCTCGATGGAGCGATAGTCAAGTTCCCGGGGCCGTGGACGCG 275
88 AlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAlaG 104
|||||
276 GCACCTATATCTCTGTGTGTGTGATCCAGATGCCCTAGCAGACAGA 325
104 uProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGly 121
|||||
326 ACCCAGACAGAGATCTGAGACATTTGGCTGTACACAGATATCAAGGCG 375
121 laAspLeuLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGln 137
|||||
376 CGGACCTGAAGAAAGGAGATTCAGGCGCAGAGTTATCAGCCTACCAG 425
138 AlaProSerProAlaHisSerGlyPheHisArg 149
|||||
426 GCTCCTCTCCCAAGCGCACAGTGGCATCGATGATAA 461

seq_name: gb_est25:A1816715

seq_documentation_block: 547 bp mRNA EST 21-DEC-1999
LOCUS A1816715
DEFINITION wj43b08.x1 NCI-CGAP-Lu19 Homo sapiens cDNA clone IMAGE:2405559 3'
similar to SW:D2_ONCVO P54187 D2 PROTEIN ;, mRNA sequence.
ACCESSION A1816715
VERSION A1816715.1 GI:5435794
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 903 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 453.
Location/Qualifiers
1. 547

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2405559"
/clone_lib="NCI-CGAP-Lu19"
/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 131 c 160 g 154 t
ORIGIN

alignment_scores:

Quality: 718.00 Length: 141
Ratio: 5.129 Gaps: 0
Percent Similarity: 99.291 Percent Identity: 98.582

alignment_block:

US-09-215-435-225 x A1816715/rev ..

Align seg 1/1 to reverse of: A1816715 from: 1 to: 547

87 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 103
|||||
547 GCGCAACCTATATCTCTGTGTGTGATCCAGATGCCCTAGCAGAGC 498
103 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysG 120
|||||
497 AGAACCCACAGACAGATCTGGAGACATTTGGCTGGTAACAGATATCAAG 448
120 lyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyr 136
|||||
447 GCGCCGACCTGAGGAGGAGAGATTCAGGCCAGGAGTTATCAGCCTAC 398
137 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPheP 153
|||||
397 CAGGCTCCCTCCCAACCGGCACACAGTGGCTTCCATCGCTACCAGTTCTT 348
153 eValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnL 170
347 TGCTATCTATCTCAGGAAGAAAAGTCATCTCTCTCTCCCAAGGAAACA 298
170 ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu 186
297 AAATCTCAGGCTCTTGGAAATGGACAGATTTTGAACCGCTTCCACTTG 248
187 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspse 203
247 GCGGAACCTGAAGCAGCACCAGTTTCATGACCCAGAACTACCAGGACTC 198

```

203 rProthrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysA 220
|||||
197 ACCAACCCCTCCAGAGTCCAGAGGAGGAGCCAGCCAGCCAGCCAGCAA. A 149
|||||
220 snGlnAlaGluLeuAlaLacys 227
|||||
148 ACCAGCGGAGATAGTGGCTGC 126
|||||

seq_name: gb_est25:AI857902

seq_documentation_block:
LOCUS AI857902 547 bp mRNA EST 21-DEC-1999
DEFINITION wj68a01.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407944 3'
similar to SW:D2_ONCVO P54187 D2 PROTEIN ;, mRNA sequence.
ACCESSION AI857902
VERSION AI857902.1 GI:5511518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 547)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA library prepared by: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
CDNA Library Arrayed by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 752 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 450.
Location/Qualifiers
1. 547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2407944"
/clone_lib="NCI-CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 102 a 130 c 160 g 155 t
ORIGIN

alignment_scores:
Quality: 718.00 Length: 141
Ratio: 5.129 Gaps: 0
Percent Similarity: 99.291 Percent Identity: 98.582

alignment_block:
US-09-215-435-225 x AI857902/rev ..
Align seg 1/1 to reverse of: AI857902 from: 1 to: 547
..

```

```

87 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 103
|||||
547 GGGCAACCTTATATCTGGGTGATGGGTGATCCAGATGCCCTTAGCAGAGC 498
|||||
103 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysG 120
|||||
497 AGAACCCAGACAGAGATTCCTGGAGACATGGCTGGTAAACAGATATCAAGG 448
|||||
120 lyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyr 136
|||||
447 GCGCCGACCTGAAGGAAGGAAGATTTCAGGGCCAGGAGTTATCAGGCTAC 398
|||||
137 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePh 153
|||||
397 CAGGCTCCCTCCCCACCGGCACACAGTGGCTTCATCGCTACCAAGTTCTT 348
|||||
153 eValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnL 170
|||||
347 TGTCTATCTTCAGGAAGGAAAGTCATCTCTCTCTCCCAAGGAAACA 298
|||||
170 ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu 186
|||||
297 AAACCTCGAGGCTCTGGGAAATGGACAGATTTCGAAACCGCTTCCACCTG 248
|||||
187 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSe 203
|||||
247 GCGAACCTGAAGCAAGCACCAGTTCATGACCAGAACTACCAAGGACTC 198
|||||
203 rProThrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysA 220
|||||
197 ACCAACCCCTCCAGGCTCCAGAGGAGGAGCCAGCCAGCCAGCCAGCAA. A 149
|||||
220 snGlnAlaGluLeuAlaLacys 227
|||||
148 ACCAGCGGAGATAGTGGCTGC 126
|||||

seq_name: gb_est69:BE044451

seq_documentation_block:
LOCUS BE044451 523 bp mRNA EST 08-JUN-2000
DEFINITION ho45g08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:3040382 3' similar to SW:D1_ONCVO P54186 D1 PROTEIN ;, mRNA
sequence.
ACCESSION BE044451
VERSION BE044451.1 GI:8361504
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1. 523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3040382"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell

```


NCI_CGAP_CCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 96 a 128 c 159 g 140 t

alignment_scores: Quality: 709.00 Length: 141 Ratio: 5.101 Gaps: 0 Percent Similarity: 98.582 Percent Identity: 97.163

alignment_block: US-09-215-435-225 x BE044451/rev .. Align seg 1/1 to reverse of: BE044451 from: 1 to: 523

87 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 103
|||||
523 GCGCAACCTATATCTGGCTGTGGTGGATCCAGATGCCCTAGCAGAC 474
103 aGluProArgGlnArgPheTyrArgHisTyrLeuValThrAspIleLysG 120
|||||
473 AGAACCCAGACAGAGATTCGGAGACATTTGGCTGGTAACAGATATCAAGG 424
120 lYalaAspLeuLysLysGlyLysIleGlnGlnGlnLeuSerAlaTyr 136
|||||
423 CGCGCGACCTGAAGGAAGGAAGATTCAGGGCCAGAGTTATCAGCCTAC 374
137 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePh 153
|||||
373 CAGGCTCCCTCCGCCACCGGCACACAGTGGCTTCATCGCTACCAAGTTCT 324
153 eValTyrLeuGlnGlnLysValIleSerLeuLeuProLysGluAsnL 170
323 TGCTATCTTCAGGAAGAAAGTATCTCTCTCTCCCAAGGAAACAA 274
170 ysThrArgGlySerTyrLysMetAspArgPheLeuAsnArgPheHisLeu 186
|||||
273 AAATCGAGGCTCTTGGAATATGGACAGATTTCTGAACCGCTTCCACCTG 224
187 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSe 203
|||||
223 GCGCAACCTGAAGCAAGCACCAGTTTCATGACCCAGAACTACCAAGGACTC 174
203 rProThrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysA 220
173 ACAACCTTCAGGCTCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 125
220 snGlnAlaGluIleAlaLaCys 227
124 ACCAGCGGAGATAGTGCTGCTGC 102

seq_name: gb_est93:BF846059

seq_documentation_block: 462 bp mRNA EST 16-JAN-2001
LOCUS BF846059
DEFINITION OVO-EN0055-181000-439-e09 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF846059
VERSION BF846059.1 GI:12233209
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 462)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

TITLE
JOURNAL
MEDLINE
COMMENT

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0055-
181000-439-e09&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 462.

FEATURES
source

1. 462
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 113 a 121 c 134 g 94 t
ORIGIN

alignment_scores: Quality: 705.00 Length: 145 Ratio: 4.896 Gaps: 0 Percent Similarity: 99.310 Percent Identity: 97.241

alignment_block:

US-09-215-435-225 x BF846059

Align seg 1/1 to: BF846059 from: 1 to: 462

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
28 ATGAGGCTGGTCACAGCAGCACTGTTACT.GGTCTCATGATGGTGTAC 76
21 rGlyAspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspG 38
|||||
77 TGGAGACGAGGATGACAGACGCCCGTGTGCCCATGAGGCCCTCTTGACG 126
38 luAspThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGly 54
|||||
127 AGGACACCCCTCTTTGCCA.GGCCTTGAAGTCTTCTACCCAGAGTTGGG 175
55 AsnIleGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysI 71
|||||
176 AACATTGGCTGCAAGGTTGTCTCTGATTGTAACTACAGACAGAAGAT 225
71 eThrSerTyrMetGluProIleValLysPheProGlyAlaValAspGlyA 88
|||||
226 CACCTCTGGATGGAGCCCATAGTCAAGTTCCCGGGGGCGGACGGCG 275
88 laThrTyrIleLeuValMetValAspProAspAlaProSerArgAlaGlu 104
|||||
276 CAACCTATATCTGGTGGTGGATCCAGATGCCCTTAGCAGCAGAA 325

Thu Nov 15 10:52:11 2001

```

105 ProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAl 121
|||||
326 CCCAGACAGATTCTGGAGACATTGGCTGGTTAACAGATATCAAGGGCGC 375
|||||

121 aAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnA 138
|||||
376 CGACCTGCAGAAAGGAGAGATTTCAGGGCCAGGAGTTATCAGCCTACCCAGG 425
|||||

138 laProSerProProAlaHisSerGlyPheHisArg 149
|||||
426 CTCCTCCCCACCGCGCACAGTGGCATCGATAAA 460
|||||

```

OM of: US-09-215-435-225_COPY_1_22 to: GenEmbl:* out_format : pfs
Date: Nov 15, 2001 12:51 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet+p2n.model -DRV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2129/app_query.fasta_1.1519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FCGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=6
-MAXLEN=2000000000 -USER=US09215435 -CGN1_1_11006 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_1_22

Query length: 22

Database: GenEmbl:*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 7698.940000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_pat2:AX083412	+	109.00	325.24	7.0e-10	681	AX083412 Sequence 104 from Pat
gb_pat2:AX083422	+	109.00	322.98	9.4e-10	897	AX083422 Sequence 114 from Pat
gb_pat1:AX060293	+	109.00	322.67	9.8e-10	932	AX060293 Sequence 1 from Pat
gb_hg12:AC037441	-	109.00	280.29	2.2e-07	162795	AC037441 Homo sapiens clone
gb_hg17:AC084081	-	109.00	279.95	2.3e-07	169645	AC084081 Homo sapiens chromo
gb_pr2:AC005723	-	58.00	122.37	140.36	127766	AC005723 Homo sapiens chromo
gb_hg1:AC008437	-	58.00	120.04	189.24	169700	AC008437 Homo sapiens chromo
gb_hg1:AC008588	-	58.00	118.44	232.33	206211	AC008588 Homo sapiens chromo
gb_r01:MUSNCPDIN	-	55.00	148.99	4.62	1585	M80840 M.musculus necln mRNA,
gb_r01:MUSNCPD	-	55.00	141.60	11.91	3899	D76440 Mouse gene for necln,
gb_hg11:AC026388_0	-	55.00	114.19	400.62	110000	AC026388 Mus musculus chromo
gb_r01:AC027298	+	55.00	109.22	757.55	201467	AC027298 Mus musculus 7 BAC
gb_ov:ORZHSCT0	+	54.00	132.94	10.03	2261	D13669 Oryzias latipes mRNA fo
gb_p12:AF193060	+	54.00	148.39	17.98	3937	AF193060 Phanerochaete chrysos
gb_hg4:AC014412	+	54.00	113.03	464.74	86446	AC014412 Drosophila melanogas
gb_hg13:AC055859	-	54.00	108.22	861.09	155295	AC055859 Homo sapiens chromo
gb_pr2:AC007370	+	54.00	107.17	984.86	176426	AC007370 Homo sapiens, clone
gb_in1:AC009213	-	54.00	106.90	1.0e+03	182525	AC009213 Drosophila melanoga
gb_in2:AE003765	-	54.00	104.68	1.4e+03	239171	AE003765 Drosophila melanoga
gb_p14:AF002844	-	53.00	104.60	1.4e+03	164839	AF002844 Oryza sativa genom
gb_p14:VCS5185	-	52.00	136.03	24.33	2444	X51616 V.carteri RNA for sulfa
gb_r01:MMU41741	+	52.00	123.92	115.04	10690	U41741 Mus musculus transcrip
gb_p13:BRBF125	+	51.00	146.24	6.57	481	L31937 Brassica rapa (clone b1
gb_b3:SFU46133	+	51.00	134.94	27.98	1905	U46133 Shigella flexneri amin
gb_ba1:AE005356	+	51.00	122.12	144.88	9083	AE005356 Escherichia coli O15
gb_ba1:AE000246	+	51.00	117.30	268.80	16338	AE000246 Escherichia coli K12
gb_ba2:D90790	-	51.00	116.50	297.89	18013	D90790 E.coli genomic DNA, KC
gb_ba2:D90791	-	51.00	115.67	331.34	19929	D90791 E.coli genomic DNA, KC
gb_pr9:HSJ245M18	+	51.00	102.05	1.9e+03	104770	AL078584 Human DNA sequence
gb_h8:HS222E13	-	51.00	100.71	2.3e+03	123360	D93241 Human DNA sequence fr
gb_hg25:HS919811	-	51.00	98.53	3.0e+03	160725	AL031302 Homo sapiens chromo
gb_hg20:AL157892	+	51.00	97.64	3.3e+03	179116	AL157892 Homo sapiens chromo
gb_hg22:ALJ512428	-	51.00	96.99	3.6e+03	193884	ALJ51248 Homo sapiens chromo
gb_hg22:ALJ392187	-	51.00	95.83	4.2e+03	223335	ALJ392187 Homo sapiens chromo
gb_ba2:AF002557	-	51.00	93.79	5.5e+03	286485	AF002557 Escherichia coli O1
gb_ba3:PWAPRNAS	+	50.50	135.06	27.54	1550	D45167 Pyrococcus sp. gene for
gb_v12:PACAPSID	+	50.50	131.28	44.75	2459	Y15938 Porcine astrovirus gene
gb_v11:AB037272	+	50.50	130.42	49.97	2730	AB037272 Porcine astrovirus ge
gb_r01:AB00071651	+	50.00	129.34	57.41	2573	AB000716 Rattus rattus gene fo
gb_p14:SPCC74	+	50.00	107.76	913.39	35637	AL031543 S.pombe chromosome I
gb_p11:AC005936	+	50.00	101.05	2.2e+03	80735	AC005936 Arabidopsis thaliana

gb_htg23:AL590149 + 50.00 94.98 4.7e+03 169181 ! AL590149 Danio rerio clon
gb_htg3:AC011676 - 50.00 93.75 5.5e+03 196574 ! AC011676 Homo sapiens clo
gb_htg2:AC009711 - 50.00 93.49 5.7e+03 202865 ! AC009711 Homo sapiens chr
gb_r02:RATADRECLB + 49.50 129.63 55.26 2050 ! L28752 Rat adrenergic recep
seq_name: gb_pat2:AX083412
seq_documentation_block:
LOCUS AX083412 681 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 104 from Patent WO0112660.
ACCESSION AX083412
VERSION AX083412.1 GI:13185249
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kato,S. and Kimura,T.
TITLE Human proteins having hydrophobic domains and dnas encoding these
proteins
JOURNAL Patent: WO 0112660-A 104 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
FEATURES
Location/Qualifiers
source 1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 179 a 186 c 183 g 133 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_1_22 x AX083412 ..
Align seg 1/1 to: AX083412 from: 1 to: 681
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
1 ATGGGTGGACAATGAGGTGGTCACAGCAGCACTGTACTGGTCTCAT 50
17 tMetValValThrGly 22
|||||
51 GATGGTGGTCACTCGA 66

seq_name: gb_pat2:AX083422
seq_documentation_block:
LOCUS AX083422 897 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 114 from Patent WO0112660.
ACCESSION AX083422
VERSION AX083422.1 GI:13185262
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kato,S. and Kimura,T.
TITLE Human proteins having hydrophobic domains and dnas encoding these
proteins
JOURNAL Patent: WO 0112660-A 114 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
FEATURES
Location/Qualifiers
source 1..897
/organism="Homo sapiens"
/db_xref="taxon:9606"
99..782
CDS
/note="unnamed protein product"

Thu Nov 15 10:52:11 2001

```

/codon_start=1
/protein_id="CAC33305.1"
/db_xref="GI:13185263"
/translation="MGWMTMLVTAALLGLMMVVTGDEDSNCPAHEALLDDTLFCQ
GLEVPFELNIGCKVVPDNNYRKITSWMEPIVKFPGAVDGTATILVMVDPAPSR
AEPQPRWRHLVTDIKGADLKKGIQGOELSAIQAPSPPAHSGFHRHYOFFVYLQEGK
VISLLPKENTRGSKWMDRLNRFHLGPEASTQFMNTQYQDSPTLQAPRERASEPKH
KNQAEIAAC"
BASE COUNT      226 a   253 c   234 g   184 t
ORIGIN

```

```

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-215-435-225_COPY_1_22 x AX083422
..

```

```

Align seg 1/1 to: AX083422 from: 1 to: 897

```

```

1 MetClyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
99 ATGGGTTGGACATGAGCGTGTACACGACGACACTGTTACTGGGTCTCAT 148
|||||
17 tMetValValThrGly 22
|||||
149 GATGGTGTCTACTGGA 164

```

```

seq_name: gb_pat1:AX060293

```

```

seq_documentation_block:
LOCUS      AX060293      932 bp      DNA      PAT      22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078802.
ACCESSION  AX060293
VERSION     AX060293.1  GI:12405782
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 932)
AUTHORS    Shimkett, R.A., Fernandes, E., Vernet, C., Yang, M., Boldog, F.L. and
Herrmann, J.L.
TITLE      Secreted polypeptides and corresponding polynucleotides
JOURNAL    Patent: WO 0078802-A 1 28-DEC-2000;
Curagen Corporation (US)
FEATURES
  source    1..932
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
  CDS       113..796
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC24866.1"
            /db_xref="GI:12405783"
            /translation="MGWMTMLVTAALLGLMMVVTGDEDSNCPAHEALLDDTLFCQ
            GLEVPFELNIGCKVVPDNNYRKITSWMEPIVKFPGAVDGTATILVMVDPAPSR
            AEPQPRWRHLVTDIKGADLKKGIQGOELSAIQAPSPPAHSGFHRHYOFFVYLQEGK
            VISLLPKENTRGSKWMDRLNRFHLGPEASTQFMNTQYQDSPTLQAPRERASEPKH
            KNQAEIAAC"

```

```

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-215-435-225_COPY_1_22 x AX060293
..

```

```

Align seg 1/1 to: AX060293 from: 1 to: 932

```

```

1 MetClyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
113 ATGGGTTGGACATGAGCGTGTACACGACGACACTGTTACTGGGTCTCAT 162
|||||
17 tMetValValThrGly 22
|||||
163 GATGGTGTCTACTGGA 178

```

```

seq_name: gb_htgl2:AC037441

```

```

seq_documentation_block:
LOCUS      AC037441      162795 bp      DNA      HTG      03-MAR-2001
DEFINITION Homo sapiens clone RP11-87E22, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
ACCESSION  AC037441
VERSION     AC037441.3  GI:13184260
KEYWORDS    HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 162795)
AUTHORS    Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens chromosome, clone RP11-87E22
JOURNAL
REFERENCE
AUTHORS

```

```

2 (bases 1 to 162795)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lenocky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Piere, N.,
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severly, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

```

```

Direct Submission
Submitted (08-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:8705171.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L9203
Center clone name: 87_E22
-----
Summary Statistics
Sequencing vector: M13; M77815; 81% of reads
Sequencing method: n/a; 19% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157025 bases at least Q40
Consensus quality: 160134 bases at least Q30
Consensus quality: 161093 bases at least Q20

```

Insert size: 158000; agarose-fp
Insert size: 161795; sum-of-ctgns
Quality coverage: 5.4 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 ba

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 15243: contig of 15243 bp in length
* 15244 15343: gap of 100 bp
* 15344 18897: contig of 3554 bp in length
* 18898 18997: gap of 100 bp
* 18998 25211: contig of 6214 bp in length
* 25212 25311: gap of 100 bp
* 25312 34838: contig of 9527 bp in length
* 34839 34938: gap of 100 bp
* 34939 45351: contig of 10413 bp in length
* 45352 45451: gap of 100 bp
* 45452 57259: contig of 11808 bp in length
* 57260 57359: gap of 100 bp
* 57360 67990: contig of 10631 bp in length
* 67991 68090: gap of 100 bp
* 68091 83399: contig of 15309 bp in length
* 83400 83499: gap of 100 bp
* 83500 105395: contig of 21896 bp in length
* 105396 105495: gap of 100 bp
* 105496 129959: contig of 24464 bp in length
* 129960 130059: gap of 100 bp
* 130060 162795: contig of 32736 bp in length.

FEATURES
Source

Location/Qualifiers
1..162795
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-87E22"
/clone_lib="RPC1-11 Human Male BAC"

1..15243
/note="assembly_fragment
clone_end:SP6
vector_side:left"
15344..18897
/note="assembly_fragment"
18998..25211
/note="assembly_fragment"
25312..34838
/note="assembly_fragment"
34939..45351
/note="assembly_fragment"
45452..57259
/note="assembly_fragment"
57360..67990
/note="assembly_fragment"
68091..83399
/note="assembly_fragment"
83500..105395
/note="assembly_fragment"
105496..129959
/note="assembly_fragment"
130060..162795
/note="assembly_fragment
clone_end:T7

BASE COUNT 42087 a 39347 c 38304 g 42057 t 1000 others
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_1_22 x AC037441/rev ..

Align seg 1/1 to reverse of: AC037441 from: 1 to: 162795

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
159302 ATGGGTTGACAAATGAGGCTGGTCACAGCAGCACTGTACTGGGTCAT 159253

17 tMetValValThrGly 22
|||||

159252 GATGGTGGTCACTGGA 159237

seq_name: gb_hlg17:AC084081

seq_documentation_block:

LOCUS AC084081 169645 bp DNA HTG 11-MAR-2001
DEFINITION Homo sapiens chromosome 8 clone RP11-89M8 map 8, WORKING DRAFT
SEQUENCE, 17 unordered pieces.

ACCESSION AC084081

VERSION AC084081.3 GI:13273380

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 169645)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 8, clone RP11-89M8

Unpublished

2 (bases 1 to 169645)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bada,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,R., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (12-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 11, 2001 this sequence version replaced gi:12583824.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11307

Center Clone name: 89_M.8

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 162624 bases at least Q40

Consensus quality: 165742 bases at least Q30

Consensus quality: 167046 bases at least Q20
Insert size: 169000; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 335: contig of 335 bp in length
* 336 435: gap of 100 bp
* 436 2289: contig of 1854 bp in length
* 2290 2389: gap of 100 bp
* 2390 4997: contig of 2608 bp in length
* 4998 5097: gap of 100 bp
* 5098 64677: contig of 59580 bp in length
* 64678 64777: gap of 100 bp
* 64778 68151: contig of 3374 bp in length
* 68152 68251: gap of 100 bp
* 68252 72655: contig of 4404 bp in length
* 72656 72755: gap of 100 bp
* 72756 78893: contig of 6138 bp in length
* 78894 78993: gap of 100 bp
* 78994 84664: contig of 5671 bp in length
* 84665 84764: gap of 100 bp
* 84765 90853: contig of 6089 bp in length
* 90854 90953: gap of 100 bp
* 90954 98425: contig of 7472 bp in length
* 98426 98525: gap of 100 bp
* 98526 106371: contig of 7846 bp in length
* 106372 106471: gap of 100 bp
* 106472 115673: contig of 9202 bp in length
* 115674 115773: gap of 100 bp
* 115774 127138: contig of 11365 bp in length
* 127139 127238: gap of 100 bp
* 127239 138436: contig of 11198 bp in length
* 138437 138536: gap of 100 bp
* 138537 149766: contig of 11230 bp in length
* 149767 149866: gap of 100 bp
* 149867 167340: contig of 17474 bp in length
* 167341 167440: gap of 100 bp
* 167441 169645: contig of 2205 bp in length.

FEATURES
Source Location/Qualifiers
1. .169645
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-89M8"
/clone_lib="RPC1-11 Human Male BAC"
1. .335
/note="assembly_fragment
clone_end:SP6
vector_side:left"
436. .2289
/note="assembly_fragment"
2390. .4997
/note="assembly_fragment"
5098. 64677
/note="assembly_fragment"
64778. .68151
/note="assembly_fragment"
68252. .72655
/note="assembly_fragment"
72756. .78893
/note="assembly_fragment"
78994. .84664
/note="assembly_fragment"

misc_feature 84765. .90853
/note="assembly_fragment"
misc_feature 90954. .98425
/note="assembly_fragment"
misc_feature 98526. .106371
/note="assembly_fragment"
misc_feature 106472. .115673
/note="assembly_fragment"
misc_feature 115774. .127138
/note="assembly_fragment"
misc_feature 127239. .138436
/note="assembly_fragment"
misc_feature 138537. .149766
/note="assembly_fragment"
misc_feature 149867. 167340
/note="assembly_fragment"
misc_feature 167441. .169645
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 44253 a 41074 c 39720 g 42960 t 1638 others
ORIGIN
alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_1_22 x AC084081/rev ..
Align seg 1/1 to reverse of: AC084081 from: 1 to: 169645
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
49659 ATGGGTGGCAATGAGCTGTCACAGCAGCAGCTTACTGGGTCTCAT 49610
17 tMetValValThrGly 22
|||||
49609 GATGGTGGTCACTGGA 49594
seq_name: gb_pr2:AC005723
seq_documentation_block:
LOCUS AC005723 127766 bp DNA PRI 28-SEP-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 194j15 (LBNL H137) - BAC clone
35G8 (LBNL H138) - BAC clone 194o13 (LBNL H106), complete
sequence.
ACCESSION AC005723 AC001452 AC001453 AC001454 AC001455 AC002172
VERSION AC005723.1 GI:3660463
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 127766)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Sequencing of human chromosome 5
Unpublished
2 (bases 1 to 127766)
REFERENCE
AUTHORS Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 127766)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1998) Human Genome Center, DOE Joint Genome

```

Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT      Sequence submitted by:
             DOE Joint Genome Institute.
FEATURES     Location/Qualifiers
             source
               1..12766
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /map="5p"
               /clone="194o13"
               /chromosome="5"
               /note="LBNL H106"
               708..731
               /note="(ATT)6"
               /rpt_type=tandem
               /rpt_unit=ATTT
               complement(1725..1847)
               /note="GRAIL 2 excellent exon, frame 0"
               3041..3141
               /rpt_family="Alu"
               3124..3172
               /note="(A)49"
               /rpt_type=tandem
               /rpt_unit=A
               5834..5857
               /note="(CCAT)6"
               /rpt_type=tandem
               /rpt_unit=CCAT
               complement(6739..6979)
               /rpt_family="L1"
               complement(7817..8218)
               /rpt_family="THE1"
               10688..10771
               /rpt_family="MIR"
               11403..11542
               /note="GRAIL 2 excellent exon, frame 0"
               12080..12159
               /rpt_family="MIR"
               complement(15524..16379)
               /rpt_family="MER4"
               16447..16574
               /rpt_family="MSTc"
               16638..16710
               /rpt_family="MLTa"
               16755..18167
               /rpt_family="MSTAR"
               18558..18711
               /rpt_family="MERS"
               complement(21242..21667)
               /rpt_family="MLT1"
               complement(21670..22014)
               /rpt_family="THE1"
               complement(22807..22937)
               /rpt_family="MER42"
               24739..24778
               /note="(AT)20"
               /rpt_type=tandem
               /rpt_unit=AT
               24779..24806
               /note="(AC)14"
               /rpt_type=tandem
               /rpt_unit=AC
               complement(25120..25532)
               /rpt_family="LOR1"
               complement(26165..26444)
               /rpt_family="Alu"
               26883..27169
               /rpt_family="Alu"
               27254..27474
               /rpt_family="Alu"
               complement(27668..27708)
               /note="GRAIL 2 excellent exon, frame 1"
               complement(29011..29154)

repeat_region      /rpt_family="Alu"
                    complement(29463..29767)
                    /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
                    complement(30117..30354)
repeat_region      /rpt_family="MER42"
                    complement(31144..31475)
                    /rpt_family="MER2"
                    complement(31781..31847)
                    /note="GRAIL 2 excellent exon, frame 1"
                    complement(31991..32229)
                    /rpt_family="L1"
                    34182..34480
                    /rpt_family="Alu"
                    36177..36461
                    /rpt_family="Alu"
                    complement(37934..38017)
                    /rpt_family="Alu"
                    38200..38560
                    /rpt_family="THE1"
                    39245..39468
                    /rpt_family="Alu"
                    complement(39823..40063)
                    /rpt_family="THE1"
                    complement(40073..40420)
                    /rpt_family="MSTAR"
                    complement(40425..40807)
                    /rpt_family="THE1"
                    42856..42977
                    /note="GRAIL 2 excellent exon, frame 2"
                    complement(43093..43800)
                    /rpt_family="L1"
                    43803..45685
                    /rpt_family="L1"
                    45747..47771
                    /standard_name="endogenous retrovirus"
                    /rpt_family="HERVH"
                    47837..48990
                    /standard_name="endogenous retrovirus"
                    /rpt_family="HERVH"
                    49149..49403
                    /standard_name="endogenous retrovirus"
                    /rpt_family="HERVH"
                    49518..49975
                    /standard_name="endogenous retrovirus"
                    /rpt_family="HERVH"
                    49976..50407
                    /rpt_family="LTR7"
                    52589..52941
                    /rpt_family="MIR2"
                    complement(52657..52793)
                    /rpt_family="Alu"
                    complement(53015..53080)
                    /rpt_family="L1"
                    54921..55022
                    /rpt_family="L1"
                    56827..56847
                    /note="(T)21"
                    /rpt_type=tandem
                    /rpt_unit=T
                    complement(56829..57122)
                    /rpt_family="Alu"
                    57181..57288
                    /rpt_family="L1"
                    58088..58379
                    /rpt_family="Alu"
                    complement(59600..59892)
                    /rpt_family="Alu"
                    60336..60367
                    /note="(GT)16"
                    /rpt_type=tandem
                    /rpt_unit=GT
                    complement(60803..60881)
                    /rpt_family="MIR"

```

```
repeat_region      complement(60883..61891)
repeat_region      /rpt_family="LTR8"
repeat_region      complement(62470..63175)
repeat_region      /rpt_family="MER44C"
repeat_region      complement(64427..64612)
repeat_region      /rpt_family="MIR"
repeat_region      65110..65263
repeat_region      /rpt_family="MER58B"
repeat_region      65358..65425
repeat_region      /rpt_family="Cheshire"
repeat_region      65422..65491
repeat_region      /rpt_family="MER58"
repeat_region      complement(66797..67059)
repeat_region      /rpt_family="Alu"
repeat_region      67827..67995
repeat_region      /rpt_family="MIR"
repeat_region      complement(69113..69282)
repeat_region      /rpt_family="MLT1"
repeat_region      complement(71015..71187)
repeat_region      /rpt_family="MST"
repeat_region      71296..71397
repeat_region      /rpt_family="MSTC"
repeat_region      72769..75031
repeat_region      /rpt_family="L1"
repeat_region      75043..75347
repeat_region      /rpt_family="Alu"

alignment_scores:
  Quality: 58.00      Length: 21
  Ratio: 3.222      Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 47.619

alignment_block:
US-09-215-435-225_COPY_1_22 x AC005723 ..
Align seg 1/1 to: AC005723 from: 1 to: 127766

2 GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMe 18
||||| ||| :||| ||||| :||| :||| :||| :||| :|||
76832 GGGTGGCGGATGATAGTAAATGCACGCTGTGTGTATGGCGGTGAAT 76881

18 tValValThrGly 22
||||| :|||
76882 GGTAGTGGTGGG 76894

seq_name: gb_htg1:AC008437

seq_documentation_block:
LOCUS AC008437 169700 bp DNA HTG 26-SEP-2000
DEFINITION Homo sapiens chromosome 5 clone CTC-329A22, WORKING DRAFT SEQUENCE,
17 unordered pieces.
ACCESSION AC008437
VERSION AC008437.5 GI:10305077
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 169700)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 169700)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Sep 26, 2000 this sequence version replaced gi:8575875.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----

Project Information
Center Project Name: 311537, H389
Center clone name: CIT-HSPC_329A22
-----
Summary Statistics
Consensus quality: 152054 bases at least Q40
Consensus quality: 160606 bases at least Q30
Consensus quality: 162232 bases at least Q20
Estimated insert size: 156000; pulse field gel estimation
Estimated coverage: 6.27 in Q20 bases; sum-of-contigs estimation
Quality coverage: 5.82 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 1042: contig of 1042 bp in length
* 1043 1142: gap of unknown length
* 1143 2192: contig of 1050 bp in length
* 2193 2292: gap of unknown length
* 2293 3380: contig of 1088 bp in length
* 3381 3480: gap of unknown length
* 3481 4986: contig of 1506 bp in length
* 4987 5086: gap of unknown length
* 5087 6137: contig of 1051 bp in length
* 6138 9208: gap of unknown length
* 9209 9309: contig of 2971 bp in length
* 9310 13759: gap of unknown length
* 13760 13859: contig of 4451 bp in length
* 13860 17412: contig of 3553 bp in length
* 17413 17512: gap of unknown length
* 17513 22391: contig of 4879 bp in length
* 22392 22491: gap of unknown length
* 22492 25367: contig of 2876 bp in length
* 25368 25467: gap of unknown length
* 25468 28671: contig of 3204 bp in length
* 28672 28771: gap of unknown length
* 28772 35739: contig of 6968 bp in length
* 35740 35839: gap of unknown length
* 35840 44103: contig of 8264 bp in length
* 44104 44203: gap of unknown length
* 44204 52953: contig of 8750 bp in length
* 52954 53053: gap of unknown length
* 53054 78367: contig of 25313 bp in length
* 78368 78466: gap of unknown length
* 78467 112257: contig of 33790 bp in length
* 112258 112356: gap of unknown length
* 112357 169700: contig of 57344 bp in length.

FEATURES
Location/Qualifiers
1..169700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-329A22"
BASE COUNT 52716 a 30825 c 30465 g 52092 t 3602 others
ORIGIN

alignment_scores:
  Quality: 58.00      Length: 21
  Ratio: 3.222      Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 47.619

alignment_block:
US-09-215-435-225_COPY_1_22 x AC008437/rev ..
Align seg 1/1 to reverse of: AC008437 from: 1 to: 169700
```


xenobiotic responsive elements (XRE) and a GC-rich region."
 /citation=[3]
 /citation=[4]
 /function="postmitotic neuron-specific expression"
 /evidence=experimental
 871..2478
 mRNA
 871..2478
 5'UTR
 CDS
 957..1934
 /codon_start=1
 /product="neclin"
 /protein_id="BAAL1183.1"
 /db_xref="GI:1040691"
 /translation="MSEQSKDLSDPNFAAPCEMDSDAVPVGIPPPASILAANLAG
 PPKAPEGMMAQQAPPEPERIEDVDPKIQQAEEGRHQPSQSPARTIPAPPAQAL
 VQKAHELMWYLVKQDKRMVFWFDMVKYKWKMSILRSTSVILARVFGHLRL
 TLNLHTEFALVKALSPEELDRVALNNRMPMTGLLILSLIYKRGAREGAVWNL
 RILGLRPKKHSTFGDVRKLTITEFVQOQNLKYQRPHTIEPEYEFWGSRRANREITK
 MOIMEFLARVFKKDPQAWPSRYREALQARALREANLAQAQPRSSVSD"
 1006
 /note="This variation gives rise to amino acid
 substitution (AA17: Ala to Val)"
 /citation=[2]
 /replace="t"
 1935..2478
 2478
 3'UTR
 polyA_site
 BASE COUNT 1107 a 895 c 892 g 1005 t
 ORIGIN
 alignment_scores:
 Quality: 55.00 Length: 16
 Ratio: 4.231 Gaps: 0
 Percent Similarity: 81.250 Percent Identity: 68.750
 alignment_block:
 US-09-215-435-225_COPY1_22 x MUSNCP/rev ..
 Align seg 1/1 to reverse of: MUSNCP from: 1 to: 3899
 2 GlycerolThrMetArgLeuValThrAlaLeuLeuGlyLeuMet 17
 |||||
 1226 GCTGAGACTCTGGCGCTGGCGCGCCCTCTCTGCGCGCTGCTG 1179
 seq_name: gb_htg11:AC026388_0
 seq_documentation_block:
 WPCOMMENT
 Sequence split into 5 fragments LOCUS AC026388 Accession AC026388
 Fragment Name Begin End
 AC026388_0 1 110000
 AC026388_1 100001 210000
 AC026388_2 200001 310000
 AC026388_3 300001 410000
 AC026388_4 400001 431552
 LOCUS AC026388 431552 bp DNA HTG 07-JAN-2001
 DEFINITION Mus musculus chromosome 7 clone RP23-426B15, WORKING DRAFT
 AC026388 SEQUENCE, 28 unordered pieces.
 AC026388.16 GI:12025600
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 431552)
 Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
 Dederich, D., Thomas, S., Okunou, G., Carroll, C., Garner, T.,
 Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,
 Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, C., Chen, Z.,
 Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,
 Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,


```
complement(34577. .35032)
/notes="region similar to Mm#S176249 vb56h02.r1 Mus
```

```
repeat_region 49133..50140
repeat_region /rpt_family="FLR13D"
repeat_region 54195..54661
repeat_region /rpt_family="L1_MM"
repeat_region complement(56789..58214)
repeat_region /rpt_family="L1VL2"
repeat_region 58666..58786
repeat_region /rpt_family="HERVL"
repeat_region 59692..60261
repeat_region /rpt_family="WT-INTERNAL"
repeat_region 60314..60343
repeat_region /rpt_family="(TTTA)n"
repeat_region complement(60438..62610)
repeat_region /rpt_family="Lx2"
repeat_region 65129..65223
repeat_region /rpt_family="RMER20"
repeat_region 65340..65610
repeat_region /rpt_family="Lx5"
repeat_region complement(65954..66079)
repeat_region /rpt_family="RMER20"
repeat_region 66080..66418
repeat_region /rpt_family="Lx5"
repeat_region 66419..66458
repeat_region /rpt_family="(CA)n"
repeat_region 66459..66836
repeat_region /rpt_family="Lx5"
repeat_region 66899..67131
repeat_region /rpt_family="Lx7"
repeat_region 67148..67201
repeat_region /rpt_family="(TG)n"
repeat_region complement(67264..67400)
repeat_region /rpt_family="Lx"
repeat_region complement(67402..67684)
repeat_region /rpt_family="ORR1A2"
repeat_region 67794..67965
repeat_region /rpt_family="B2"
repeat_region 68188..68440
repeat_region /rpt_family="Lx9"
misc_feature complement(68713..69166)

alignment_scores:
  Quality: 55.00 Length: 16
  Ratio: 4.231 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-09-215-435-225_COPY_1_22 x AC027298 ..
Align seg 1/1 to: AC027298 from: 1 to: 201467
2 GlyTTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMet 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196746 GCGTGACGCTGGGCTGTGGCGGCGCCCTCTGCGGCTGCTG 196793

seq_name: gb_ov:ORZHSC70

seq_documentation_block: 2261 bp mRNA VRT 13-APR-1999
LOCUS ORZHSC70
DEFINITION Oryzias latipes mRNA for heat shock protein 70 cognate, complete cds.
ACCESSION D13669
VERSION D13669.1 GI:538526
KEYWORDS heat shock protein 70 cognate.
SOURCE Oryzias latipes (strain:HB32) fibroblast-like cell_line:cultured Medaka cell OL32 cDNA to mRNA.
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 2261)
AUTHORS Arai,A., Naruse,K., Mitani,H. and Shima,A.
TITLE Cloning and characterization of cDNAs for 70-kDa heat-shock proteins (Hsp70) from two fish species of the genus Oryzias

JOURNAL Jpn. J. Genet. 70 (3), 423-433 (1995)
MEDLINE 95358880
REFERENCE 2 (bases 1 to 2261)
AUTHORS Naruse,K.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1992) to the DDBJ/EMBL/GenBank databases. Kiyoshi Naruse, School of Science, University of Tokyo., Department of Biological Sciences; Hongo 7-3-1, Bunkyo-Ku, Tokyo 113, Japan (E-mail:naruse@biol.s.u-tokyo.ac.jp, Tel:03-3812-2111(ex.4443), Fax:03-3816-1965)

FEATURES
    Location/Qualifiers
        1..2261
            /organism="Oryzias latipes"
            /strain="HB32"
            /db_xref="taxon:8090"
            /cell_line="cultured Medaka cell OL32"
            /cell_type="fibroblast-like"
            122..2182
                /codon_start=1
                /product="heat shock protein 70 cognate"
                /protein_id="BAA76887.1"
                /db_xref="GI:4589737"
                /translation="MSKGPAYGIDLTGYSCVGFQHGKVEIIANDQGNRTTPSYVAF
                TDERLIGDAAKNOVAMNPTNTVFDAKRLIGRRFDHVQSDMNDWPNFVINDNTRPK
                VOVEYKGETSKFYPEEYSSMLTKMKETAAYLGKTVNNNAVITVPAYFNDSORQATKD
                AGTISGLNVLRIINEPTAAATAYGLDKKVGSENVLLFDLGGGTFDVSILITIEGIFE
                VKSTAGDTHLGGEDFDNRWNHFIATPKRYKKDISDNKRAVRRLRSACERAKRTLSS
                STQASIEDSLYEGVDFTYSITRAFELNADLFGLDPVEKSLRDAMDKQQLHDI
                VLVGSGTRIPKIQKLQDFNGKELNKSINPEAVAYGAQAAILSGDKSENVDLL
                LLDVTPLSLGIETAGVMTVLIKRNTIPTKTQTFYSDNQFGLNQVYEVGAMTK
                DNLLGKFLTGIPALWCPOIEVTFDIDANGINNVSADKSTGKENKITINDKGRLL
                SKEDIERWQEAERYKAEDDVORDKVSANKLESYAFNMKSTVDEKLAGISDEDKO
                KILDKNEVISWLDKNQTAERDEYVHQKELEKVCNPITIKLYSAGSGMPGCGQRECO
                EVSLELVLLAVALLDQFLPKLLIKHSMSPSKIFTTKTLFKCNMQSINKSVMLNQFPF
                ITIT"
BASE COUNT 647 a 496 c 538 g 580 t
ORIGIN

alignment_scores:
  Quality: 54.00 Length: 20
  Ratio: 3.176 Gaps: 0
Percent Similarity: 85.000 Percent Identity: 50.000

alignment_block:
US-09-215-435-225_COPY_1_22 x ORZHSC70 ..
Align seg 1/1 to: ORZHSC70 from: 1 to: 2261
1 MetGlyTTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeu 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
669 ATGGCGTGACAAAAGTTGGCTCCGAGAGGAATGTTCTTATCTTGTAT 718

seq_name: gb_pl2:AF193060

seq_documentation_block: 3937 bp DNA PLN 12-NOV-2000
LOCUS AF193060
DEFINITION Phanerochaete chrysosporium NADPH-dependent cytochrome P450 oxidoreductase (CPR) gene, complete cds.
ACCESSION AF193060
VERSION AF193060.1 GI:11138368
KEYWORDS Phanerochaete chrysosporium.
SOURCE Phanerochaete chrysosporium.
ORGANISM Phanerochaete chrysosporium.
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Hymenomycetidae; Aphyllophorales; Corticiaceae; Phanerochaete.
REFERENCE 1 (bases 1 to 3937)
AUTHORS Yadav,J.S. and Loper,J.C.
```

```

alignment_scores:      Length: 16
Quality: 54.00
Ratio: 4.154
Gaps: 0
Percent Similarity: 81.250    Percent Identity: 68.750

alignment_block:
US-09-215-435-225_COPY_1_22_x_AF193060    ..
Align seg 1/1    to: AF193060    from: 1    to: 3937

2    GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuClvLeuMet 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707    GCGTGGACTTTCGAAACGAAACCGCTCGCTCTCGTTTGGT 754

seq_name: gb_htg4:AC014412

seq_documentation_block:
LOCUS    AC014412    86446 bp    DNA
DEFINITION    Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION    AC014412
VERSION    AC014412.1    GI:5436923
KEYWORDS    HTG; HTGS_PHASE2.

```


THIS PAGE BLANK (USPTO)

OM of: US-09-215-435-225_COPY_1_22 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:29 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-FGAPOP=4.500 -FGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@Cgnl_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXP
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_1_22

Query length: 22

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA40582 +	109.00	303.87	5.9e-09	447	
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480 +	109.00	299.97	9.7e-09	681	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA397660 +	109.00	298.18	1.2e-08	826	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA300013 +	109.00	298.18	1.2e-08	826	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA242252 +	109.00	297.49	1.3e-08	890	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA963350 +	109.00	297.42	1.4e-08	897	
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94490 +	109.00	297.36	1.4e-08	903	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA15582 +	109.00	297.30	1.4e-08	909	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA377539 +	109.00	297.06	1.4e-08	932	
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA384882 +	109.00	296.87	1.5e-08	952	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA78431 +	109.00	296.16	1.6e-08	1028	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA78426 +	49.00	93.40	3.1e+03	29729	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA13175 -	48.00	124.31	59.29	777	
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAV44217 +	48.00	98.85	1.6e+03	12127	
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AAV52281 +	47.00	126.29	46.00	461	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA31710 +	47.00	125.78	49.10	487	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA34773 +	47.00	125.35	51.87	510	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA345669 +	47.00	125.32	52.11	512	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA32882 +	47.00	120.68	94.53	845	
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AA728519 -	47.00	49.50	6.6e+05	183012	
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AA72063 -	46.50	91.46	4.0e+03	16941	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA234737 +	46.00	122.77	72.24	495	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA78455 +	46.00	121.76	82.24	552	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA78544 +	46.00	104.49	753.89	3561	
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA71597 +	46.00	103.93	809.31	3780	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA239542 +	45.50	102.25	1.0e+03	3886	
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AAV74399 +	45.00	121.77	82.13	405	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA398102 -	45.00	119.56	109.03	519	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA32772 +	45.00	118.95	117.90	544	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA395279 +	45.00	115.88	174.90	765	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA398101 +	45.00	113.42	239.18	839	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA252569 +	45.00	115.06	195.46	993	
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AA355661 +	45.00	113.46	238.46	993	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA385781 +	45.00	107.70	499.23	1849	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA398035 -	45.00	78.71	2.0e+04	42235	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA398043 +	44.50	108.87	429.79	1397	
/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ88155 +	44.50	104.95	709.92	2131	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ88162 +	44.00	126.63	44.02	176	
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAV89602 +	44.00	125.31	52.15	203	
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA374745 +	44.00	120.53	96.27	340	

/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA48965 + 44.00 117.13 149.00 491
/cgnl_9/gcgdata/geneseq/geneseq/NA1996.DAT:AA12701 + 44.00 114.19 217.12 674
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ80075 + 44.00 110.52 347.83 100
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AA33249 + 44.00 107.42 517.61 140
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AA01349 + 44.00 106.96 548.96 147

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AA40582

seq_documentation_block:

ID AAX40582 standard; cDNA: 447 BP.

XX AAX40582;

DT 18-JUN-1999 (first entry)

DE Human secreted protein 5' EST SEQ ID No: 182.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX Homo sapiens.

OS W09906550-A2.

PN 11-FEB-1999.

PD 31-JUL-1998; 98WO-IB01232.

XX 01-AUG-1997; 97US-0905144.

XX (GBST) GENSET.

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

DR WPI: 1999-153780/13.

DR P-PSDB; AAY11860.

PT New isolated prostate-derived nucleic acids - used to develop products which may have cytokine, immune regulatory, haematopoiesis regulating, anti-inflammatory or tumour inhibition activity

PS Claim 1; Page 285; 675pp; English.

XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for human secreted proteins expressed in prostate, and encode the proteins given in AAY11716 to AAY11993 respectively. The proteins given represent the signal peptide and an N-terminal fragment of a secreted protein. The nucleic acid sequences can be used for producing secreted human gene products. They can also be used to develop products for diagnosis and therapy. The proteins obtained may have cytokine activity, cell proliferation and differentiation activity, haematopoiesis regulating activity, tissue growth regulating activity, reproductive hormone regulating activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, tumour inhibition activity or other activities. The products can be used in forensic, gene therapy and chromosome mapping procedures. The sequences can also be used for obtaining corresponding promoter sequences. The nucleic acids encoding the signal peptides can be used for directing extracellular secretion of a polypeptide or the insertion of a polypeptide into a membrane, or importing a polypeptide into a cell.

SQ Sequence 447 BP; 98 A; 110 C; 136 G; 99 T; 4 other;

alignment_scores:

Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
CC wound healing, as nutritional sources or supplements e.g. as amino acid,
CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
CC processing and utilisation of dietary fat, protein, carbohydrate,
CC vitamins and minerals, to effect behavioural characteristics, to affect
CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein.
XX
SQ Sequence 681 BP; 179 A; 186 C; 183 G; 133 T; 0 other;

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x AAF94480 ..
Align seg 1/1 to: AAF94480 from: 1 to: 681
1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
1 ATGGGTGGACAATGAGGTGGTGCACAGCAGCACTGTTACTGGGTCTCAT 50
17 tMetValValThrGly 22
|||||
51 GATGGTGGTCACTGGA 66

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AAF97660

seq_documentation_block:
ID AAX97660 standard; DNA: 826 BP.

XX AC AAX97660;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein coding sequence, SEQ ID NO. 124.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.
XX
OS Homo sapiens.
XX
PN W0931236-A2.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB02122.
XX
PR 10-AUG-1998; 98US-0096116.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI: 1999-385906/32.
XX
DR P-PSDB; AAY35976.
XX
PT New isolated human secreted proteins
XX
PS Claim 1; Page 255; 516pp; English.

CC alignment_block:
US-09-215-435-225_COPY_1_22 x AAX40582 ..
Align seg 1/1 to: AAX40582 from: 1 to: 447
1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
85 ATGGGTGGACAATGAGGTGGTGCACAGCAGCACTGTTACTGGGTCTCAT 134
17 tMetValValThrGly 22
|||||
135 GATGGTGGTCACTGGA 150

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480

seq_documentation_block:
ID AAF94480 standard; CDNA: 681 BP.

XX AC AAF94480;
XX
DT 04-JUN-2001 (first entry)
XX
DE Human hydrophobic domain containing protein clone HP03880 cDNA #104.
XX
KW Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
KW cytostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response; ss.
XX
OS Homo sapiens.
XX
PN W0200112660-A2.
XX
PD 22-FEB-2001.
XX
PF 10-AUG-2000; 2000WO-JP05356.
XX
PR 17-AUG-1999; 99JP-0230344.
PR 07-SEP-1999; 99JP-0252551.
PR 01-OCT-1999; 99JP-0281132.
PR 22-OCT-1999; 99JP-0301624.
PR 04-NOV-1999; 99JP-0313877.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
XX
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR WPI: 2001-160059/16.
DR P-PSDB; AAB88590.

XX Human proteins with hydrophobic domains and the DNAs which encode them
XX are useful for treating autoimmune disorders, burns and tumors and for
XX screening novel pharmaceuticals -
XX
XX Claim 3; Page 426-427; 518pp; English.
XX
XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
XX AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
XX anti-HIV, neuroprotective, antianemic, vulnery, antiulcer,
XX osteopathic, anti-inflammatory and cytostatic activities, and can be
XX used in gene therapy. (I) can be used as pharmaceuticals and as antigens
XX to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
XX probes for genetic diagnosis and gene sources for gene therapy or for
XX producing (I) in large quantities. Cells containing (II) are used for
XX the detection of ligands or receptors corresponding to membrane or
XX secretory proteins and to screen small molecule novel pharmaceuticals.
XX Antibodies directed to (I) can be used for the detection, quantification
XX and purification of (I). Activities of (I) may include cytokine and cell
XX proliferation/differentiation function, immune stimulating or suppressing
XX activity, haematopoiesis regulating activity, tissue growth activity,

XX This sequence represents an extended human secreted protein coding
 CC sequence of the invention. The secreted proteins can be used in treating
 CC or controlling a variety of human conditions. The secreted proteins may
 CC act as cytokines or may affect cellular proliferation or differentiation
 CC or may act as immune system regulators, haematopoiesis regulators, tissue
 CC growth regulators, regulators of reproductive hormones or cell movement
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
 CC tumour inhibition activity. The DNAs can be used in forensic procedures
 CC to identify individuals or in diagnostic procedures to identify
 CC individuals having genetic diseases resulting from abnormal expression of
 CC the genes corresponding to the extended cDNAs. They are also useful for
 CC constructing a high resolution map of the human chromosomes. They can
 CC also be used for gene therapy to control or treat genetic diseases.
 XX
 SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAX97660 ..

Align seg 1/1 to: AAX97660 from: 1 to: 826

```

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTTGGACAAATGAGGCTGCTCAGCAGCAGCTGTACTGGGTCTCAT 64

17 tMetValValThrGly 22
|||||
65 GATGGTGGTCACTGGA 80

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC00013

seq_documentation_block:

ID AAC00013 standard; cDNA; 826 BP.

XX AAC00013;

DT 06-OCT-2000 (first entry)

DE Human secreted protein cDNA sequence #4.

XX Human; secreted protein; 5' EST; expressed sequence tag; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 15..698

FT /*tag= a

FT /product= "secreted protein"

PN EP1033401-A2.

XX 06-SEP-2000.

PD 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GSET) GENSET.

PA Dumas Milne Edwards J, Duclert A, Giordano J;

PI WPI; 2000-500381/45.

XX P-PSDB; AAG00016.

DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Example 19; SEQ ID 7; 71pp + CD-ROM; English.

XX The present sequence is a full length cDNA encoding a human
 CC secreted protein. The cDNA was obtained from a 5' EST using first and
 CC second strand synthesis procedures. 5' ESTs were prepared from total
 CC human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region
 CC (UTR) of the mRNA because they are often obtained from oligo-dT primed
 CC cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors.

SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAC00013 ..

Align seg 1/1 to: AAC00013 from: 1 to: 826

```

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
15 ATGGGTTGGACAAATGAGGCTGCTCAGCAGCAGCTGTACTGGGTCTCAT 64

17 tMetValValThrGly 22
|||||
65 GATGGTGGTCACTGGA 80

```

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AA242252

seq_documentation_block:

ID AA242252 standard; cDNA; 826 BP.

XX AA242252;

DT 01-FEB-2000 (first entry)

DE Human phosphatidylethanolamine-binding protein encoding cDNA.

XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification; ss.

OS Homo sapiens.

XX WO9953051-A2.

XX 21-OCT-1999.

XX 09-APR-1999; 99WO-IB00712.

XX 09-APR-1998; 98US-0057719.

XX 28-APR-1998; 98US-0069047.

XX (GSET) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-038446/03.

DR

DR P-PSDB; AAY64647.
 XX Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX
 XX
 PS Example 21; Page 168-169; 837pp; English.
 XX
 CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY65438 represent the EST-related proteins corresponding to AA242265 to
 CC AA243052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-215-435-225_COPY_1_22 x AA242252 ..

Align seg 1/1 to: AA242252 from: 1 to: 826

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 |||||
 15 ATGGTGTGGACAAATGAGCTGTGTACACACACACTGTACTGGGTCTCAT 64
 |||||
 17 MetValValThrGly 22
 |||||
 65 GATGGTGGTCACTGGA 80

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA96350

seq_documentation_block:

ID AAA96350 standard; cDNA; 890 BP.

XX AAA96350;

AC (first entry)

DE cDNA encoding a novel polypeptide designated PRO4408.

XX Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;
 KW PRO1889; PRO1890; PRO1887; PRO1785; PRO4353; PRO4357; PRO4405; PRO4356;
 KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;
 KW PRO4424; PRO4422; PRO4430; PRO4499; tumour; obesity; diabetes;
 KW insulinemia; kidney disorder; Bergers disease; nephropathy;
 KW Schonein-Henoch purpura; celiac disease; dermatitis herpetiformis;
 KW Crohns disease; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 89..760

FT /*tag= a

FT sig_peptide 89..154
 FT /*tag= b

PN WO200056889-A2.

XX 28-SEP-2000.

XX 01-MAR-2000; 2000WO-US05601.

XX 23-MAR-1999; 99US-0125774.

XX 23-MAR-1999; 99US-0125778.

XX 24-MAR-1999; 99US-0125826.

XX 31-MAR-1999; 99US-0127035.

XX 05-APR-1999; 99US-0127706.

XX 21-APR-1999; 99US-0130359.

XX 27-APR-1999; 99US-0131270.

XX 27-APR-1999; 99US-0131272.

XX 27-APR-1999; 99US-0131291.

XX 04-MAY-1999; 99US-0132371.

XX 04-MAY-1999; 99US-0132379.

XX 04-MAY-1999; 99US-0132383.

XX 25-MAY-1999; 99US-0135750.

XX 08-JUN-1999; 99US-0138166.

XX 20-JUL-1999; 99US-0144791.

XX 03-AUG-1999; 99US-0146970.

XX 09-DEC-1999; 99US-0170262.

XX (GETH) GENENTECH INC.

XX Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
 PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;

XX WPI: 2000-628263/60.

XX P-PSDB; AAB18923.

XX Novel secreted and transmembrane polypeptides useful for diagnosing
 PT tumour in a mammal, for identifying agonists and antagonists of the
 PT polypeptide and for therapeutic use

XX Claim 2; Fig 29; 222pp; English.

XX The present sequence encodes a secreted or transmembrane polypeptide.
 CC The specification describes polypeptides designated PRO1484, PRO4334,
 CC PRO1122, PRO1889, PRO1890, PRO1887, PRO1785, PRO4353, PRO4357, PRO4405,
 CC PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,
 CC PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is
 CC useful for diagnosing tumour in a mammal. The polypeptides, their
 CC agonists and antagonists are useful treating a condition associated with
 CC expression or activity of the polypeptide. Conditions treated include
 CC obesity, diabetes or hyper-or hypo-insulinemia. The polypeptides are
 CC capable of inducing proliferation of mammalian kidney mesangial cells
 CC and are therefore useful for treating kidney disorders associated with
 CC decreased mesangial cell function such as Bergers disease or other
 CC nephropathies associated with Schonein-Henoch purpura, celiac disease,
 CC dermatitis herpetiformis or Crohns disease. The nucleic acids may be used
 CC to generate transgenic animals for use in development and screening of
 CC therapeutically useful reagents and also for chromosome identification
 CC and tissue typing.

XX Sequence 890 BP; 228 A; 246 C; 234 G; 182 T; 0 other;

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_1_22 x AAA96350

Align seg 1/1 to: AAA96350 from: 1 to: 890

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17

|||||
89 ATGGGTGGCAATGAGGTGGTCACAGCAGCACTGTACTGGGTCTCAT 138
17 tMetValValThrGly 22
|||||
139 GATGGTGGTCACTGGA 154

seq_name: /cgn1_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF94490

seq_documentation_block:
ID AAF94490 standard; cDNA; 897 BP.
XX
AC AAF94490;
XX
DT 04-JUN-2001 (first entry)
DE Human hydrophobic domain containing protein clone HP03880 cDNA #114.
XX
KW Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
KW cystostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response; ss.
XX
OS Homo sapiens.
XX
PN WO200112660-A2.
XX
PD 22-FEB-2001.
XX
PF 10-AUG-2000; 2000WO-JP05356.
XX
PR 17-AUG-1999; 99JP-0230344.
PR 07-SEP-1999; 99JP-0252551.
PR 01-OCT-1999; 99JP-0281132.
PR 22-OCT-1999; 99JP-0301624.
PR 04-NOV-1999; 99JP-0313877.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
WPI: 2001-160059/16.
P-PSDB; AAB88590.

Human proteins with hydrophobic domains and the DNAs which encode them are useful for treating autoimmune disorders, burns and tumors and for screening novel pharmaceuticals -
Claim 4; Page 442-444; 518pp; English.
CC AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
CC anti-HIV, neuroprotective, antianemic, vulnery, antiulcer,
CC osteopathic, anti-inflammatory and cystostatic activities, and can be
CC used in gene therapy. (I) can be used as pharmaceuticals and as antigens
CC to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
CC probes for genetic diagnosis and gene sources for gene therapy or for
CC producing (I) in large quantities. Cells containing (II) are used for
CC the detection of ligands or receptors corresponding to membrane or
CC secretory proteins and to screen small molecule novel pharmaceuticals.
CC Antibodies directed to (I) can be used for the detection, quantification
CC and purification of (I). Activities of (I) may include cytokine and cell
CC proliferation/differentiation function, immune stimulating or suppressing
CC activity, haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
CC wound healing, as nutritional sources or supplements e.g. as amino acid,

CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
CC processing and utilisation of dietary fat, protein, carbohydrate,
CC vitamins and minerals, to effect behavioural characteristics, to affect
CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein.
XX
SQ Sequence 897 BP; 226 A; 253 C; 234 G; 184 T; 0 other;

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_1_22 x AAF94490 ..
Align seg 1/1 to: AAF94490 from: 1 to: 897
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
99 ATGGTGGACATGAGCTGGTCACAGCAGCACTGTACTGGGTCTCAT 148
17 tMetValValThrGly 22
|||||
149 GATGGTGGTCACTGGA 164

seq_name: /cgn1_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA15582

seq_documentation_block:
ID AAA15582 standard; cDNA; 903 BP.
XX

AC AAA15582;
XX

DT 01-AUG-2000 (first entry)
XX

DE Human phospholipid binding protein 2, PLBP2 gene.
XX

KW Human; phospholipid binding protein; PLBP2; foetal development disorder;
KW reproduction disorder; cell proliferation disorder; immune response;
KW autoimmune disorder; AIDS; infertility; cytostatic; immunosuppressive;
KW gene therapy; hereditary neuropathy;
KW phosphatidylethanolamine binding protein D1; PE-BP D1; ss.
XX

OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 88..771
FT /*tag= a
FT /product= "Human PLBP2"

XX
XX US6063767-A.
XX

PD 16-MAY-2000.
XX

PF 09-DEC-1998; 98US-0208718.
XX

PR 28-OCT-1997; 97US-0958820.
XX

PA (INCY-) INCYTE PHARM INC.
XX

PI Corley NC, Shah P, Lal P, Hillman JL;
XX

DR WPI: 2000-375529/32.
XX

DR P-PSDB; AAY94263.
XX

XX New purified phospholipid binding proteins 1 and 2 useful for
PT diagnosing, treating or preventing diseases disorders associated with
PT fetal development, reproduction, cell proliferation, and the immune
PT response -
XX

XX Example 5; Fig 2; 37pp; English.
XX

Thu Nov 15 10:52:11 2001

CC The present sequence is the phospholipid binding protein 2 (PLBP2) gene.
 CC This gene is expressed in lung, prostate and heart tissues. Also, the
 CC protein is expressed in foetal tumour tissues. PLBP2 may be used for the
 CC diagnosis, prevention, or treatment of disorders associated with foetal
 CC development (e.g. hereditary neuropathies), reproduction (e.g.
 CC infertility), cell proliferation (e.g. cancers), and the immune response
 CC (AIDS). PLBP2 antibodies may also be developed for potential drug
 CC screening or to quantitate PLBP2 gene expression in biopsied tissues. The
 CC PLBP2 gene may be administered for gene therapy of disorders associated
 CC with PLBP2. PLBP2 has high homology with the phosphatidylethanolamine
 CC binding protein PL, PE-BP DI, of *Onchocerca volvulus*. PE-BP DI is thought
 CC to play a role in transport or signal mechanisms between membranes and
 CC the cytoplasm.

XX Sequence 903 BP; 222 A; 251 C; 242 G; 188 T; 0 other;
 SQ

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAA15582 ..

Align seg 1/1 to: AAA15582 from: 1 to: 903

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 |||||
 88 ATGGGTTGGCAATGAGCTGTCACACGACACTGTTACTGGGTCAT 137
 |||||
 17 tMetValValThrGly 22
 |||||
 138 GATGGTGTCACCTGGA 153

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC77539

seq_documentation_block:

ID AAC77539 standard; cDNA; 909 BP.

AC AAC77539;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF3094 polynucleotide sequence SEQ ID NO:6187.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antilnfammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antilnfammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.
 PA Shimkets RA, Leach M;
 PI WPI; 2000-602362/57.
 XX P-PSDB; AAB43330.
 DR Novel nucleic acids and peptides derived from open reading frame X,
 XX useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PT Claim 5; Page 5369-5370; 5507pp; English.
 XX

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antilnfammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX protein in gene therapy
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, diabetes mellitus,
 CC graft vs host disease, cardiovascular disease, cholesterol ester storage, systemic lupus
 CC hypertension, hypothyroidism, cholesterol ester storage, SCID, AIDS, viral,
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antilnfammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 909 BP; 232 A; 254 C; 236 G; 186 T; 1 other;
 SQ

alignment_scores:
 Quality: 109.00 Length: 22
 Ratio: 4.955 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAC77539 ..

Align seg 1/1 to: AAC77539 from: 1 to: 909

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 |||||
 100 ATGGGTTGGCAATGAGCTGTCACACGACACTGTTACTGGGTCAT 149
 |||||
 17 tMetValValThrGly 22
 |||||
 150 GATGGTGTCACCTGGA 165

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC84882

seq_documentation_block:

ID AAC84882 standard; cDNA; 932 BP.

XX AAC84882;

XX 20-APR-2001 (first entry)

XX Human SEC1 nucleic acid sequence (clone ID 3445452).

XX SECX; cytostatic; gynecological; gene therapy; screening assay; human;
 KW chromosomal mapping; forensic biology; cell proliferation; cancer;
 KW cell differentiation; immune associated disorder; gestational disease;
 KW SEC1; ss.

```
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 113..781
XX CDS /*tag= a
XX FT /product= "SECL"
XX FT 113..181
XX FT /*tag= b
XX FT 182..778
XX FT /*tag= c
XX
XX PN W0200078802-A2.
XX
XX PD 28-DEC-2000.
XX
XX PF 23-JUN-2000; 2000WO-US17328.
XX
XX PR 23-JUN-1999; 99US-0140584.
XX PR 20-JUL-1999; 99US-0144722.
XX PR 16-SEP-1999; 99US-0154520.
XX PR 22-JUN-2000; 2000US-0604286.
XX
XX PA (CURA-) CURAGEN CORP.
XX
XX PI Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
XX PI Herrmann JL;
XX
XX DR WPI: 2001-071385/08.
XX DR P-PSDB; AAB48368.
XX
XX PT Polynucleotides encoding SECX proteins useful for treating disease
XX PT characterized by an aberrant level of cell proliferation and/or
XX PT differentiation like cancer or immune associated disorders -
XX
XX PS Claim 3; Fig 1; 132pp; English.
XX
XX CC The invention relates to human SECX polypeptides and polynucleotides
XX CC encoding them. The SECX polypeptides can be expressed by standard
XX CC recombinant methodology. The SECX polypeptides are useful for treating
XX CC or preventing a SECX-associated disorder. The invention is useful in
XX CC screening assays; detection assays (e.g. chromosomal mapping, cell and
XX CC tissue typing, forensic biology); predictive medicine (diagnostic assays,
XX CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
XX CC methods of treatment (e.g. therapeutic and prophylactic), especially
XX CC disorders characterized by aberrant cell proliferation and/or
XX CC differentiation like cancer or immune associated disorders or gestational
XX CC disease. The present sequence represents a SECL nucleic acid sequence.
XX
XX SQ Sequence 932 BP; 244 A; 257 C; 244 G; 187 T; 0 other;

alignment_scores:
    Quality: 109.00 Length: 22
    Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x AAC84882 ..

Align seg 1/1 to: AAC84882 from: 1 to: 932

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
113 ATGGGTTGCACATGAGCTGGTCACAGCAGCACTGTTACTGGTCTCAT 162
|||||
17 tMetValValThrGly 22
|||||
163 GATGGTGGTCACTGGA 178

seq_name: /cgni_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA78431

seq_documentation_block:
ID AAA78431 standard; cDNA: 952 BP.
```

```
XX AAA78431;
XX
XX AC 20-NOV-2000 (first entry)
XX
XX DT Human secreted protein gene 46 SEQ ID NO:61.
XX
XX DE Human; secreted protein; cytostatic; antianaemic; antidiabetic;
XX KW antiinflammatory; ophthalmologic; antirheumatic; antiarthritic;
XX KW antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
XX KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
XX KW immune system disorder; angiogenesis; hyperproliferative disorder;
XX KW cardiovascular disorder; apoptosis; neurological disease;
XX KW infectious disease; wound healing; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200035937-A1.
XX
XX PD 22-JUN-2000.
XX
XX PF 16-DEC-1999; 99WO-US29950.
XX
XX PR 17-DEC-1998; 98US-0112809.
XX PR 18-DEC-1998; 98US-0113006.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
XX PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX
XX DR WPI: 2000-431566/37.
XX DR P-PSDB; AAB24487.
XX
XX PT Forty seven human nucleic acids encoding secreted proteins, useful in
XX PT the treatment, prevention and diagnosis of cancers, disorders of the
XX PT immune system, angiogenesis disorders, neurological diseases and
XX PT hyperproliferative disorders -
XX
XX PS Claim 1; Page 478; 562pp; English.
XX
XX CC The polynucleotide sequence given in AAA78381 to AAA78432 encode the
XX CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
XX CC proteins have activities based on the tissues and cells the genes are
XX CC expressed in. Examples of activities include: cytostatic; antianaemic;
XX CC antidiabetic; antiinflammatory; ophthalmologic; antirheumatic;
XX CC antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
XX CC nootropic; neuroprotective; antimicrobial and antiparkinsonian.
XX CC Human secreted protein polynucleotides, polypeptides, antagonists and/or
XX CC agonists may be useful in treating, preventing, and/or diagnosing other
XX CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
XX CC disorders of the immune system; (c) angiogenesis disorders; (d)
XX CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
XX CC associated with increase apoptosis; (g) neurological diseases; and
XX CC (h) infectious diseases. They are also used to promote wound healing.
XX CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
XX CC exemplification of the present invention.
XX
XX SQ Sequence 952 BP; 310 A; 247 C; 223 G; 172 T; 0 other;
```

```
alignment_scores:
    Quality: 109.00 Length: 22
    Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x AAA78431 ..

Align seg 1/1 to: AAA78431 from: 1 to: 952

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
|||||
|||||
```

63 ATGGTTGACAAATGAGGTGGTGCACAGCAGCACTGTTACTGGGTCCTCAT 112
Ratio: 4.955
Percent Similarity: 100.000 Percent Identity: 100.000

17 tMetValValThrGly 22
113 GATGGTGCTCACTGGA 128

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA78426

seq_documentation_block:
ID AAA78426 standard; cDNA; 1028 BP.
XX AC AAA78426;
XX DT 20-NOV-2000 (first entry)
XX DE Human secreted protein gene 46 SEQ ID NO:56.
XX KW Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW antipsoriatic; angiogenic; cardiant; anti-HIV; nootropic;
KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW immune system disorder; angogenesis; hyperproliferative disorder;
KW cardiovascular disorder; apoptosis; neurological disease;
KW infectious disease; wound healing; ss.
XX OS Homo sapiens.
XX PN WO200035937-A1.
XX PD 22-JUN-2000.
XX PF 16-DEC-1999; 99WO-US29950.
XX PR 17-DEC-1998; 98US-0112809.
XX PR 18-DEC-1998; 98US-0113006.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX WP: 2000-431566/37.
XX P-PSDB; AAB24482.

Forty seven human nucleic acids encoding secreted proteins, useful in the treatment, prevention and diagnosis of cancers, disorders of the immune system, angogenesis disorders, neurological diseases and hyperproliferative disorders -
Claim 1; Page 475; 562pp; English.

The polynucleotide sequence given in AAA78381 to AAA78432 encode the human secreted proteins given in AAB24437 to AAB24604. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic; antianaemic; antidiabetic; antiinflammatory; ophthalmological; antirheumatic; antiarthritic; antipsoriatic; antimicrobial and antiparkinsonian. Human secreted protein polynucleotides, polypeptides, antagonists and/or agonists may be useful in treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions such as: (a) cancers; (b) disorders of the immune system; (c) angogenesis disorders; (d) hyperproliferative disorders; (e) cardiovascular diseases; (f) diseases associated with increase apoptosis; (g) neurological diseases; and (h) infectious diseases. They are also used to promote wound healing.
AAA78372 to AAA78380 and AAB24436 represent sequences used in the exemplification of the present invention.

Sequence 1028 BP; 309 A; 268 C; 259 G; 186 T; 6 other;

63 ATGGTTGACAAATGAGGTGGTGCACAGCAGCACTGTTACTGGGTCCTCAT 112
Ratio: 4.955
Percent Similarity: 100.000 Percent Identity: 100.000

17 tMetValValThrGly 22
113 GATGGTGCTCACTGGA 128

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA78426

seq_documentation_block:
ID AAA78426 standard; cDNA; 1028 BP.
XX AC AAA78426;
XX DT 20-NOV-2000 (first entry)
XX DE Human secreted protein gene 46 SEQ ID NO:56.
XX KW Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW antipsoriatic; angiogenic; cardiant; anti-HIV; nootropic;
KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW immune system disorder; angogenesis; hyperproliferative disorder;
KW cardiovascular disorder; apoptosis; neurological disease;
KW infectious disease; wound healing; ss.
XX OS Homo sapiens.
XX PN WO200035937-A1.
XX PD 22-JUN-2000.
XX PF 16-DEC-1999; 99WO-US29950.
XX PR 17-DEC-1998; 98US-0112809.
XX PR 18-DEC-1998; 98US-0113006.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX WP: 2000-431566/37.
XX P-PSDB; AAB24482.

Forty seven human nucleic acids encoding secreted proteins, useful in the treatment, prevention and diagnosis of cancers, disorders of the immune system, angogenesis disorders, neurological diseases and hyperproliferative disorders -
Claim 1; Page 475; 562pp; English.

The polynucleotide sequence given in AAA78381 to AAA78432 encode the human secreted proteins given in AAB24437 to AAB24604. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic; antianaemic; antidiabetic; antiinflammatory; ophthalmological; antirheumatic; antiarthritic; antipsoriatic; antimicrobial and antiparkinsonian. Human secreted protein polynucleotides, polypeptides, antagonists and/or agonists may be useful in treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions such as: (a) cancers; (b) disorders of the immune system; (c) angogenesis disorders; (d) hyperproliferative disorders; (e) cardiovascular diseases; (f) diseases associated with increase apoptosis; (g) neurological diseases; and (h) infectious diseases. They are also used to promote wound healing.
AAA78372 to AAA78380 and AAB24436 represent sequences used in the exemplification of the present invention.

Sequence 1028 BP; 309 A; 268 C; 259 G; 186 T; 6 other;

alignment_scores: Quality: 49.00 Length: 18
 Ratio: 3.267 Gaps: 0
 Percent Similarity: 83.333 Percent Identity: 55.556

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAX13175/rev ..
 Align seg 1/1 to reverse of: AAX13175 from: 1 to: 29729

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
 10490 TTAATGCTTTTACCAGGCGTGTAGTCATGGGACTATGATGCCAGTAAC 10441
 21 rGly 22
 10440 TGGT 10437

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV44217
 seq_documentation_block:
 ID AAV44217 standard; DNA; 777 BP.
 XX AC AAV44217;
 XX DT 19-JUL-1999 (first entry)
 XX DE Lettuce resistance gene AC15-2B.
 XX KW Resistance gene; AC15-2B; RG gene; lettuce; disease resistance;
 KW pest resistance; virus; fungus; protozoan; bacterium; nematode;
 KW crop protection; transgenic plant; fingerprinting; ss.
 OS Lactuca sativa.
 XX WO9830083-A1.
 PN 16-JUL-1998.
 PD 09-JAN-1998; 98WO-US00615.
 PF 10-JAN-1997; 97US-0781734.
 PR (REGC) UNIV CALIFORNIA.
 XX PA Meyers B, Michelmore RW, Shen K;
 XX PI WPI; 1998-398692/34.
 XX DR New resistance gene nucleic acids - useful to produce disease
 PT resistant plants, e.g. lettuce, Lactuca sativa, and to detect
 PT resistance genes, e.g. to fingerprint cultivars
 XX PS Disclosure; Page 93-97; 183pp; English.

XX This is the nucleotide sequence of the lettuce AC15-2B gene
 CC which hybridises to lettuce RG2 resistance genes (see AAV44199-214)
 CC of the invention. The invention relates to families of resistance
 CC genes, termed RG or RLG, from lettuce. RG families RG1-RG5 and RG7
 CC have been identified, each RG family being defined as a group of
 CC polypeptide sequences that have at least 60% amino acid sequence
 CC identity. Individual members of an RG family typically map to the
 CC same genomic locus. RG nucleic acid sequences (see AAV44188-257) can
 CC be used to confer resistance in plants to a variety of pests,
 CC including viruses, fungi, nematodes, insects and bacteria.
 CC Sequences from within the RG genes can be used to fingerprint
 CC cultivars or germplasm for the presence of desired resistance
 CC genes. Promoters of RG genes can be used to drive heterologous
 CC gene expression under conditions in which RG genes are expressed.
 CC RG polynucleotides can also be used to produce antisense and
 CC ribozyme molecules useful for inhibiting RG activity in plants.
 CC The invention also provides RG proteins (see AAV64150-98) that are
 CC encoded by these RG genes.

XX SQ Sequence 777 BP; 261 A; 148 C; 120 G; 246 T; 2 other;
 alignment_scores: Quality: 48.00 Length: 19
 Ratio: 3.200 Gaps: 0
 Percent Similarity: 78.947 Percent Identity: 42.105

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAV44217/rev ..
 Align seg 1/1 to reverse of: AAV44217 from: 1 to: 777

3 TrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetVa 19
 577 TGGTTATTCGGAGTGTACCAATTCCTACTTGTGGACTCTCACTCTC 528
 19 lValThr 21
 527 TATCTCA 521

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV52281
 seq_documentation_block:
 ID AAV52281 standard; DNA; 12127 BP.
 XX AC AAV52281;
 XX DT 23-OCT-1998 (first entry)
 XX DE Streptococcus pneumoniae genome fragment SEQ ID NO:148.
 XX KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
 KW computer readable medium; vaccine; pharmaceutical composition; ds.
 XX OS Streptococcus pneumoniae.
 XX PN WO9818931-A2.
 XX PD 07-MAY-1998.
 XX PF 30-OCT-1997; 97WO-US19588.
 XX PR 31-OCT-1996; 96US-0029960.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
 PI Kunsch CA, Rosen CA;
 XX WPI; 1998-272225/24.
 XX Computer-readable medium with recorded Streptococcus pneumoniae
 PT polynucleotide sequences - useful in diagnostic kits and assays, and
 PT pharmaceutical compositions and vaccines for Streptococcus
 PT pneumoniae
 XX Claim 1; Page 991-998; 1409pp; English.

XX The present invention describes a computer readable medium which has
 CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
 CC recorded on it, or a representative fragment or a sequence at least 95%
 CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
 CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
 CC Streptococcus pneumoniae. The present invention also describes an
 CC isolated nucleic acid molecule encoding a homologue of any of the
 CC fragments of the S pneumoniae genome (SEQ ID NO:1 to 391) where the
 CC nucleic acid molecule is produced by a process comprising: (a) screening
 CC a genomic DNA library using as a probe a target sequence defined by any
 CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
 CC library which contain sequences that hybridise to the target sequence and
 CC isolating the nucleic acid molecules from the members; or (b) isolating

Thu Nov 15 10:52:11 2001

CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
 CC molecules whose nucleotide sequence is homologous to amplification
 CC primers derived from the fragment of the S. pneumoniae genome to prime
 CC the amplification and isolating the amplified sequences. The computer
 CC readable medium can be used in a computer-based system for identifying
 CC fragments of the S. pneumoniae genome of commercial importance, or
 CC expression modulating fragments of the S. pneumoniae genome. Products
 CC from the present invention can be used in diagnosis kits and assays, and
 CC pharmaceutical compositions and vaccines for S. pneumoniae.
 XX
 SQ Sequence 12127 BP; 3743 A; 2512 C; 2345 G; 3527 T; 0 other;

alignment_scores:
 Quality: 48.00 Length: 21
 Ratio: 2.824 Gaps: 0
 Percent Similarity: 80.952 Percent Identity: 47.619

alignment_block:
 US-09-215-435-225_COPY_1_22 x AAV52281 ..

Align seg 1/1 to: AAV52281 from: 1 to: 12127

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
 :::||| ||| :::::|||||:::|||||:::|||||:::|||||
 1479 TTAGGAATGGCATGTGCTTAACCGCTCGTTACTATTAGGATAT 1528
 17 tMetValValThr 21
 |:::| |
 1529 GATTTGCCAACC 1541

OM of: US-09-215-435-225_COPY_1_22 to: Issued_Patents_NA:* out_format : pfs

Date: Nov 15, 2001 5:03 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000
-GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@CGNL_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_1_22

Query length: 22

Database: Issued_Patents_NA:*

Database sequences: 351203

Database length: 113238999

Search time (sec): 228.030000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4	+	109.00	318.52	4.3e-10	903
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US96-05320A-1342	+	47.00	135.16	7.05	3
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-743-6378-25	-	47.00	127.65	18.46	845
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-526-8408-25	-	47.00	127.65	18.46	845
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-198-122-1	+	44.50	111.08	154.71	2124
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-328-111-159	+	44.00	116.75	74.73	1002
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-307-896-2	-	44.00	112.77	124.47	1516
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-344-914-1	-	44.00	112.77	124.47	1516
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-11808-2	-	44.00	112.51	128.73	1558
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-054-077C-1	-	44.00	104.97	338.57	3415
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-326-568-18	+	44.00	103.61	403.05	3934
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-077-848A-1	+	44.00	103.58	404.57	3946
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-211-640-1	+	44.00	103.58	404.57	3946
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-378-536-1	+	44.00	103.58	404.57	3946
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US94-03547-1	+	44.00	103.58	404.57	3946
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-471-119A-1	-	44.00	79.80	8.5e+03	46899
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2	+	44.00	36.16	6.7e+05	4403765
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-1	+	44.00	36.15	6.7e+05	4411529
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-166-346A-4	-	43.00	125.64	23.90	288
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-171-982-1	-	43.00	123.34	32.11	366
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-091-405-1	-	43.00	112.11	135.56	1178
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US91-00909-3	+	43.00	111.60	144.69	1242
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-148-215A-3	-	43.00	109.67	185.27	1518
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-580-680-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-480-156A-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-354-961-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-581-094-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-580-665-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-581-142-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-581-103-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US94-05442A-2	-	43.00	106.33	284.50	2150
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-944-449-3	-	43.00	106.32	284.66	2150
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-027-337-9	+	43.00	103.45	411.35	2900
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-027-337-1	+	43.00	102.67	454.94	3147
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-812-7	+	43.00	98.17	809.61	5024
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-827-7	+	43.00	98.17	809.61	5024
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-521-177-7	+	43.00	98.17	809.61	5024
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-362-577C-7	+	43.00	98.17	809.61	5024
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-920-828-7	+	43.00	98.17	809.61	5024
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-149-103-1	-	42.50	100.58	594.60	3330

/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-451-883-1 - 42.50 100.58 594.60 3330
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-393-734-1 - 42.50 99.68 667.12 3656
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-894-489-1 - 42.50 99.68 667.12 3656
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-393-734-3 - 42.50 90.42 2.2e+03 9592
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-894-489-3 - 42.50 90.42 2.2e+03 9592

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4

seq_documentation_block:
; Sequence 4, Application US/09208718
; Patent No. 6063767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PHOSPHOLIPID BINDING PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0379 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGUT12
; CLONE: 3126479
; US-09-208-718-4

alignment_scores:

Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_1_22 x US-09-208-718-4 ...

Align seg 1/1 to: US-09-208-718-4 from: 1 to: 903

1 MetGlyTriThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17

|||||

88 ATGGGTGGACATGAGGTGGTCACAGCAGCAGCTGTACTGGGTCTCAT 137

17 tMetValValThrGly 22

|||||

138 GATGGTGGTCACTGGA 153

```

seq_name: /cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:seq:PCT-US96-05320A-1342
seq_documentation_block:
; Sequence 1342, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20003-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April22, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.014PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1342:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US96-05320A-1342

alignment_scores:
Quality: 47.00 Length: 18
Ratio: 3.357 Gaps: 0
Percent Similarity: 77.778 Percent Identity: 55.556

alignment_block:
US-09-215-435-225_COPY1_22 x PCT-US96-05320A-1342
Align seg 1/1 to: PCT-US96-05320A-1342 from: 1 to: 387
5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
13 ATGAAAAAATTACATTAGCTTTAGCTTCAGCTTTAGCTGTCGAC 62

seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:seq:US-08-743-637B-25
seq_documentation_block:
; Sequence 25, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Haemophilus influenzae
; US-08-743-637B-25

alignment_scores:
Quality: 47.00 Length: 18
Ratio: 3.357 Gaps: 0
Percent Similarity: 77.778 Percent Identity: 55.556

alignment_block:
US-09-215-435-225_COPY1_22 x US-08-743-637B-25/rev
Align seg 1/1 to reverse of: US-08-743-637B-25 from: 1 to: 845
5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
423 ATGAAAAAATTACATTAGCTTTAGCTTTAGCTTCAGCTTTAGCTGTCGAC 374
21 rGly 22
|||||
373 AGGT 370

```

2 GlyTprThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMe 18
|||||
1305 GGGTGGACTATTGGACAAAGTACGCC.....TTGGTATTCAGCTC 1345
18 tValValThr 21

Thu Nov 15 10:52:12 2001

us-09-215-435-225_copy_1_22.rni

```

; SEQ ID NO 2
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-08-307-896-2

alignment_scores:
  Quality: 44.00      Length: 16
  Ratio: 3.143       Gaps: 0
  Percent Similarity: 87.500   Percent Identity: 56.250

alignment_block:
  US-09-215-435-225_COPY_1_22 x US-08-307-896-2 from: 1 to: 1516
  Align seg 1/1 to reverse of: US-08-307-896-2 from: 1 to: 1516

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeu 16
  ::::: ||:::||||::: :::::||||:::||||:::
81 GTTGGCTCACGCTCGCTCTCTCGTTCGCTCGCTCGCTCTT 34

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-344-914-1

seq_documentation_block:
; Sequence 1, Application US/09344914
; Patent No. 6110664
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-ALPHA-S1 EXPRESSION
; FILE REFERENCE: RYS-0068
; CURRENT APPLICATION NUMBER: US/09/344,914
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 1
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1155)
; US-09-344-914-1

alignment_scores:
  Quality: 44.00      Length: 16
  Ratio: 3.143       Gaps: 0
  Percent Similarity: 87.500   Percent Identity: 56.250

alignment_block:
  US-09-215-435-225_COPY_1_22 x US-09-344-914-1/rev ..
  Align seg 1/1 to reverse of: US-09-344-914-1 from: 1 to: 1516

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeu 16
  ::::: ||:::||||::: :::::||||:::||||:::
81 GTTGGCTCACGCTCGCTCTCTCGTTCGCTCGCTCGCTCTT 34

seq_name: /cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-11808-2

seq_documentation_block:
; Sequence 2, Application PC/TUS9511808
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi V.
; TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
; TITLE OF INVENTION: ADENYLYL
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue and
; ADDRESSEE: Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
;

seq_documentation_block:
; Sequence 159, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endese, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1002)
; OTHER INFORMATION: n = A,T,C or G
; US-09-328-111-159

alignment_scores:
  Quality: 44.00      Length: 17
  Ratio: 3.385       Gaps: 0
  Percent Similarity: 76.471   Percent Identity: 52.941

alignment_block:
  US-09-215-435-225_COPY_1_22 x US-09-328-111-159 ..
  Align seg 1/1 to: US-09-328-111-159 from: 1 to: 1002

3 TrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetva 19
  ::::: ||:::||||||| ||:::||||| :::::||||
548 TGGAGTTGCTTTTGGTACACAGCTCCATTGCTACTCCCATTTATTGGT 597

19 1 19
598 T 598

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-307-896-2

seq_documentation_block:
; Sequence 2, Application US/08307896C
; Patent No. 6034071
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi
; TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
; TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
; FILE REFERENCE: 29770
; CURRENT APPLICATION NUMBER: US/08/307,896C
; CURRENT FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
;
;
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694314
; FILING DATE: 01-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 40937
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 406..2733
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 337..2733
; US-08-054-077C-1

alignment_scores:
      Quality: 44.00      Length: 18
      Ratio: 2.750      Gaps: 0
Percent Similarity: 88.889      Percent Identity: 44.444

alignment_block:
US-09-215-435-225_COPY_1_22 x US-08-054-077C-1/rev ..
Align seg 1/1 to reverse of: US-08-054-077C-1 from: 1 to: 3415

      2 GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetMe 18
      ||||| :|:::||||| :|:::||||| :|:::||||| :|:::||||| :|:::
1290 GGATGGATAGTCATCTCGTTGGATCGTGTACTGTGTGGCTCGTTCA 1241

      18 tval 19
      :|||
      1240 GGTG 1237

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-226-568-18

seq_documentation_block:
; Sequence 18, Application US/09226568
; Patent No. 6001992
; GENERAL INFORMATION:
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcusson, Eric G.
; TITLE OF INVENTION: Antisense Modulation of No. 6001992el Anti-apoptotic
; TITLE OF INVENTION: bcl-2-Related Proteins
; FILE REFERENCE: ISPH-0337
; CURRENT APPLICATION NUMBER: US/09/226,568
; CURRENT FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3934
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
; OTHER INFORMATION: sequence
; US-09-226-568-18

alignment_scores:
      Quality: 44.00      Length: 21

```

Ratio: 2.444 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-226-568-18 ..
Align seg 1/1 to: US-09-226-568-18 from: 1 to: 3934

1 MetGlyTrpThrMetArgLeuVal..ThrAlaAlaLeuLeuGlyLeu 16
||||| ||||| ||| :||||| |||||
2417 ATGGCTTGGAAAGCAGGCTAGTCTAACCATGGTCTATTATTAGGCTTG 2466

17 MetMetValValThrGly 22
:|||||
2467 CTTGTTACACACAGGT 2484

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-077-848A-1

seq_documentation_block:
; Sequence 1, Application US/08077848A
; Patent No. 5470955
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,848A
; FILING DATE: 16-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-2845
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note= "When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
; US-08-077-848A-1

alignment_scores:
Quality: 44.00 Length: 23
Ratio: 2.444 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-08-077-848A-1 ..

US-09-215-435-225_COPY_1_22 x US-08-077-848A-1 ..
Align seg 1/1 to: US-08-077-848A-1 from: 1 to: 3946

1 MetGlyTrpThrMetArgLeuVal..ThrAlaAlaLeuLeuGlyLeu 16
||||| ||||| ||| :||||| |||||
2417 ATGGCTTGGAAAGCAGGCTAGTCTAACCATGGTCTATTATTAGGCTTG 2466

17 MetMetValValThrGly 22
:|||||
2467 CTTGTTACACACAGGT 2484

seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-211-640-1

seq_documentation_block:
; Sequence 1, Application US/09211640
; Patent No. 6020466
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,640
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/441,375
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note= "When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
; US-09-211-640-1

alignment_scores:
Quality: 44.00 Length: 23
Ratio: 2.444 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-211-640-1 ..

Align seg 1/1 to: US-09-211-640-1 from: 1 to: 3946

```
1 MetGlyTrpThrMetArgLeuVal...ThrAlaAlaLeuLeuLeuGlyLeu 16
||||:||||| ||||||||| ||| :|||||||||
2417 ATGGCTTGGAAAGCAGGCTAGTCTAACCATGGTGTATTATTAGGCTTG 2466

17 MetMetValValThrGly 22
||||:|||||
2467 CTGTGTACACACACAGGT 2484
```

seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-378-536-1

```
seq_documentation_block:
; Sequence 1, Application US/09378536
; Patent No. 6200763
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,536
; FILING DATE:
```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,848
; FILING DATE: 16-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note= "when nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
US-09-378-536-1
```

alignment_scores:
Quality: 44.00 Length: 23
Ratio: 2.444 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-378-536-1 ..

Align seg 1/1 to: US-09-378-536-1 from: 1 to: 3946

```
1 MetGlyTrpThrMetArgLeuVal...ThrAlaAlaLeuLeuLeuGlyLeu 16
||||:||||| ||||||||| ||| :|||||||||
2417 ATGGCTTGGAAAGCAGGCTAGTCTAACCATGGTGTATTATTAGGCTTG 2466

17 MetMetValValThrGly 22
||||:|||||
2467 CTGTGTACACACACAGGT 2484
```

seq_name: /cgnl_7/ptodata/1/ina/PCTUS_COMB.seq:PCT-US94-03547-1

```
seq_documentation_block:
; Sequence 1, Application PC/TUS9403547
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: MYELOID CELL LEUKEMIA ASSOCIATED GENE
; TITLE OF INVENTION: MCL-1
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03547
; FILING DATE: 31-MAR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note= "when nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
PCT-US94-03547-1
```

alignment_scores:
Quality: 44.00 Length: 23
Ratio: 2.444 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 56.522

alignment_block:
US-09-215-435-225_COPY_1_22 x PCT-US94-03547-1 ..

Align seg 1/1 to: PCT-US94-03547-1 from: 1 to: 3946

```
1 MetGlyTrpThrMetArgLeuVal...ThrAlaAlaLeuLeuLeuGlyLeu 16
||||:||||| ||||||||| ||| :|||||||||
2417 ATGGCTTGGAAAGCAGGCTAGTCTAACCATGGTGTATTATTAGGCTTG 2466
```


[illegible]

Thu Nov 15 10:52:12 2001

us-09-215-435-225_copy_1_22.rnrm

```
; SOFTWARE: PERL Program
; SEQ ID NO 7685
; LENGTH: 223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00799304
; NAME/KEY: unsure
; LOCATION: 17
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-853-7685

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-534-853-7685 ..
Align seg 1/1 to: US-09-534-853-7685 from: 1 to: 223

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
64 ATGGGTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 113

17 tMetValValThrGly 22
|||||
114 GATGGTGGTCACTGGA 129

seq_name: /cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-035-171-1718

seq_documentation_block:
; Sequence 1718, Application US/09035171
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuve, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: LUNG TUMOR
; NUMBER OF SEQUENCES: 2026
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/035.171
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/039,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0327P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
;
```

```
1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
54 ATGGGTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 103

17 tMetValValThrGly 22
|||||
104 GATGGTGGTCACTGGA 119

seq_name: /cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-534-853-7707

seq_documentation_block:
; Sequence 7707, Application US/09534853
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuve, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND MEME
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: PD-1016 CIP
; CURRENT APPLICATION NUMBER: US/09/534,853
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
; SOFTWARE: PERL Program
; SEQ ID NO 7707
; LENGTH: 190
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00790232
US-09-534-853-7707

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-534-853-7707 ..
Align seg 1/1 to: US-09-534-853-7707 from: 1 to: 190

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
54 ATGGGTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 103

17 tMetValValThrGly 22
|||||
104 GATGGTGGTCACTGGA 119

seq_name: /cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-534-853-7685

seq_documentation_block:
; Sequence 7685, Application US/09534853
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuve, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND MEME
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: PD-1016 CIP
; CURRENT APPLICATION NUMBER: US/09/534,853
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
;
```

```
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 1718:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 2723986H1
US-09-035-171-1718

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-035-171-1718 ..
Align seg 1/1 to: US-09-035-171-1718 from: 1 to: 239

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
10 ATGGGTGGACAATGAGCGTGGTCACAGCAGCACTGTACTGGGTCTCAT 59

17 tMetValValThrGly 22
|||||
60 GATGGTGGTCACTGGA 75

seq_name: /cgnl_7/ptodata/1/pna/US095C_COMB.seq:US-09-540-212A-10721

seq_documentation_block:
; Sequence 10721, Application US/09540212A
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
; FILE REFERENCE: PD-1034 CIP
; CURRENT APPLICATION NUMBER: US/09/540,212A
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 67551
; SOFTWARE: PERL Program
; SEQ ID NO 10721
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00358906
US-09-540-212A-10721

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-540-212A-10721 ..
Align seg 1/1 to: US-09-540-212A-10721 from: 1 to: 239

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
10 ATGGGTGGACAATGAGCGTGGTCACAGCAGCACTGTACTGGGTCTCAT 59

17 tMetValValThrGly 22

; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 1718:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 2723986H1
US-09-035-171-1718

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-039-416-1718 ..
Align seg 1/1 to: US-09-039-416-1718 from: 1 to: 239

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
10 ATGGGTGGACAATGAGCGTGGTCACAGCAGCACTGTACTGGGTCTCAT 59

17 tMetValValThrGly 22
|||||
60 GATGGTGGTCACTGGA 75

seq_name: /cgnl_7/ptodata/1/pna/US088_COMB.seq:US-08-856-524-163
```

us-09-215-435-225_copy_1_22.rnrm

Thu Nov 15 10:52:12 2001

```

seq_documentation_block:
; Sequence 163, Application US/08856524
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN FETAL LUNG
; NUMBER OF SEQUENCES: 1379
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,524
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0177P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1354794
; US-08-856-524-163

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-08-856-524-163
Align seg 1/1 to: US-08-856-524-163 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
47 ATGGGTTGGACAATGAGGCTGGTGCACAGCAGCAGCTGTTACTGGGTCTCAT 96

17 tmetvalvalthrGly 22
|||||
97 GATGGTGGTCACTGGA 112

seq_name: /cgnl_7/ptodata/1/pna/US088_COMB.seq:US-08-856-524-905

seq_documentation_block:
; Sequence 1249, Application US/08862968
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.

```

```
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN PROSTATE
; NUMBER OF SEQUENCES: 1716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,968
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,396
; FILING DATE: MAY 24, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0175 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 1249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1347609
; US-08-862-968-1249

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-08-862-968-1249 ..
Align seg 1/1 to: US-08-862-968-1249 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
69 ATGGGTGGACAATGAGCTGGTCACAGCAGCACTGTTACTGGTCTCAT 118

17 tMetValValThrGly 22
|||||
119 GATGGTGGTCACTGGA 134

seq_name: /cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-534-853-7694

seq_documentation_block:
; Sequence 7694, Application US/09534853
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND M
; FILE REFERENCE: PD-1016 CIP
; CURRENT APPLICATION NUMBER: US/09/534,853
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
; SOFTWARE: PERL Program
; SEQ ID NO 7700
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00410317
; NAME/KEY: unsure
; LOCATION: 176
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-853-7700

alignment_scores:
```

```
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND M
; FILE REFERENCE: PD-1016 CIP
; CURRENT APPLICATION NUMBER: US/09/534,853
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
; SOFTWARE: PERL Program
; SEQ ID NO 7694
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00489356
; NAME/KEY: unsure
; LOCATION: 22
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-853-7694

alignment_scores:
  Quality: 109.00      Length: 22
  Ratio: 4.955        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-534-853-7694 ..
Align seg 1/1 to: US-09-534-853-7694 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
69 ATGGGTGGACAATGAGCTGGTCACAGCAGCACTGTTACTGGTCTCAT 118

17 tMetValValThrGly 22
|||||
119 GATGGTGGTCACTGGA 134

seq_name: /cgnl_7/ptodata/1/pna/US095B_COMB.seq:US-09-534-853-7700

seq_documentation_block:
; Sequence 7700, Application US/09534853
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND M
; FILE REFERENCE: PD-1016 CIP
; CURRENT APPLICATION NUMBER: US/09/534,853
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
; SOFTWARE: PERL Program
; SEQ ID NO 7700
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00410317
; NAME/KEY: unsure
; LOCATION: 176
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-853-7700

alignment_scores:
```

```

Quality: 109.00      Length: 22
Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-534-853-7700  ..
Align seg 1/1 to: US-09-534-853-7700 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
47 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 96

17 tMetValValThrGly 22
|||||
97 GATGGTGGTCACTGGA 112

seq_name: /cgnl_7/ptodata/1/pna/us095B_COMB.seq:us-09-534-853-7704

seq_documentation_block:
; Sequence 7704, Application us/09534853
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SURFACE STRUCTURE AND MEME
; FILE REFERENCE: PD-1016 CIP
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 28399
; SOFTWARE: PERL Program
; SEQ ID NO 7704
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00156135
; NAME/KEY: unsure
; LOCATION: 129, 145, 152, 157, 160, 163, 170, 175, 181, 189, 194, 196, 200, 202, 205,
; LOCATION: 218-219, 223, 225, 235, 238
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-853-7704

alignment_scores:
Quality: 109.00      Length: 22
Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-534-853-7704  ..
Align seg 1/1 to: US-09-534-853-7704 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
47 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 96

17 tMetValValThrGly 22
|||||
97 GATGGTGGTCACTGGA 112

seq_name: /cgnl_7/ptodata/1/pna/us6001_COMB.seq:us-60-017-766-163

seq_documentation_block:
; Sequence 163, Application us/60017766
; GENERAL INFORMATION:

```

```

; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 1379
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/017,766
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0177P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1354794
; US-60-017-766-163

alignment_scores:
Quality: 109.00      Length: 22
Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-60-017-766-163  ..
Align seg 1/1 to: US-60-017-766-163 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
47 ATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCTCAT 96

17 tMetValValThrGly 22
|||||
97 GATGGTGGTCACTGGA 112

seq_name: /cgnl_7/ptodata/1/pna/us6001_COMB.seq:us-60-017-766-905

seq_documentation_block:
; Sequence 905, Application us/60017766
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.

```


;; APPLICANT: Akertblom, Ingrid E.
;; APPLICANT: Delegeane, Angelo M.
;; APPLICANT: Naughton, Rebecca E.
;; APPLICANT: Klingler, Tod M.
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;; HUMAN FETAL LUNG
;; NUMBER OF SEQUENCES: 1379
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;; STREET: 3174 PORTER DRIVE
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/60/017,766
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: LUTHER, BARBARA J.
;; REGISTRATION NUMBER: 33954
;; REFERENCE/DOCKET NUMBER: PD-0177P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 855-0555
;; TELEFAX: (415) 852-0195
;; INFORMATION FOR SEQ ID NO: 905:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 241 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: 1357688
US-60-017-766-905

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-60-017-766-905 ..

Align seg 1/1 to: US-60-017-766-905 from: 1 to: 241

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
47 ATGGGTGGACAATGAGCTGGTGCACAGCAGCAGCTGTACTGGGTCTCAT 96

17 tMetValValThrGly 22
|||||
97 GATGGTGGTCACTGGA 112

seq_name: /cgnl_7/ptodata/1/pna/US088_COMB.seq:US-08-808-904-3732

seq_documentation_block:
; Sequence 3732, Application US/08808904
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akertblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.

;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;; TITLE OF INVENTION: ASTHMATIC LUNG
;; NUMBER OF SEQUENCES: 5094
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;; STREET: 3174 PORTER DRIVE
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/808,904
;; FILING DATE: HEREWITH
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/012,698
;; FILING DATE: FEBRUARY 29, 1996
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/013,365
;; FILING DATE: MARCH 13, 1996
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/027,122
;; FILING DATE: SEPTEMBER 27, 1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CERRONE, MICHAEL C.
;; REGISTRATION NUMBER: 39,132
;; REFERENCE/DOCKET NUMBER: PD-0120 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 855-0555
;; TELEFAX: (415) 845-4166
;; INFORMATION FOR SEQ ID NO: 3732:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 244 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: 1985045
US-08-808-904-3732

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-08-808-904-3732 ..

Align seg 1/1 to: US-08-808-904-3732 from: 1 to: 244

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
87 ATGGGTGGACAATGAGCTGGTGCACAGCAGCAGCTGTACTGGGTCTCAT 136

17 tMetValValThrGly 22
|||||
137 GATGGTGGTCACTGGA 152

us-09-215-435-225_copy_1_22.rnppm

Thu Nov 15 10:52:12 2001

Thu Nov 15 10:52:13 2001

us-09-215-435-225_copy_1_22.rnpn

```
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 17
||||:|||| ||| ||||||||||| :||| |||||||
449 ATGCCCTGGCTCATGGCAAGGTTACTGCAATATGATCTGGGCTTGC 498
||||:|||| ||| ||||||||||| :||| |||||||
17 tMetValValThrGly 22
||||:|||| ||| ||||||||||| :||| |||||||
499 TACATTTGGTACATGGG 514
||||:|||| ||| ||||||||||| :||| |||||||

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-4155

seq_documentation_block:
; Sequence 4155, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4155
; LENGTH: 1902
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(390)
; US-09-815-242-4155

alignment_scores:
Quality: 47.50 Length: 21
Ratio: 2.969 Gaps: 1
Percent Similarity: 76.190 Percent Identity: 57.143

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-815-242-4155 ..
Align seg 1/1 to: US-09-815-242-4155 from: 1 to: 1902

3 Trp...ThrMetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMe 18
||||:|||| ||| ||||||||||| :||| |||||||
1443 TGGCGGACGCTTCGGCTTCGACGTCGTCGCTTCTCTCGGCTACCGCG 1492
||||:|||| ||| ||||||||||| :||| |||||||
1493 AGTTGTTCGAGGT 1505
||||:|||| ||| ||||||||||| :||| |||||||

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-7131

seq_documentation_block:
; Sequence 7131, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4155
; LENGTH: 1902
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(390)
; US-09-815-242-7131

alignment_scores:
Quality: 47.00 Length: 18
Ratio: 3.357 Gaps: 0
Percent Similarity: 77.778 Percent Identity: 55.556

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-815-242-7131 ..
Align seg 1/1 to: US-09-815-242-7131 from: 1 to: 390

5 MetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetValValTh 21
||||:|||| ||| ||||||||||| :||| |||||||
13 ATGAAAAAATTAACATTAGCATTTGGTTTACGCTTTAGCTGTGAC 62
||||:|||| ||| ||||||||||| :||| |||||||
21 rGly 22
||||:|||| ||| ||||||||||| :||| |||||||
63 AGGT 66
||||:|||| ||| ||||||||||| :||| |||||||

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-4631

seq_documentation_block:
; Sequence 4631, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
```

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4631
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4631

alignment_scores:
  Quality: 45.50      Length: 19
  Ratio: 2.844       Gaps: 1
  Percent Similarity: 84.211  Percent Identity: 52.632

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-815-242-4631 ..
Align seg 1/1 to: US-09-815-242-4631 from: 1 to: 1311

      2 GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMe 18
      ||||| :: :: ||||| ::|||::|||::|||::|||::|||
      415 GGCTGG...CTAACGATTGTTATCGCGCACTCGTATTAGGATTAATCAT 461
      18 tValVal 20
      ||::|||
      462 GGCAGTC 468

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-8323

seq_documentation_block:
; Sequence 8323, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-031A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8323
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-8323

alignment_scores:
  Quality: 45.50      Length: 19
  Ratio: 2.844       Gaps: 1
  Percent Similarity: 84.211  Percent Identity: 52.632

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-815-242-8323 ..
Align seg 1/1 to: US-09-815-242-8323 from: 1 to: 1347

      2 GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMe 18
      ||||| :: :: ||||| ::|||::|||::|||::|||::|||
      433 GGCTGG...CTAACGATTGTTATCGCGCACTCGTATTAGGATTAATCAT 479
      18 tValVal 20
      ||::|||
      480 GGCAGTC 486

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-974-684-6

seq_documentation_block:
; Sequence 6, Application US/09974684
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Human Nucleic Acid Sequences Encoding
; FILE REFERENCE: G-Protein Coupled Receptor-like Homologs and Uses Therefor
; CURRENT APPLICATION NUMBER: US/09/974,684
; CURRENT FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-974-684-6

alignment_scores:
  Quality: 45.00      Length: 23
  Ratio: 2.647       Gaps: 1
  Percent Similarity: 73.913  Percent Identity: 47.826

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-974-684-6 ..
Align seg 1/1 to: US-09-974-684-6 from: 1 to: 1110

      2 GlyTrp.....ThrMetArgLeuValThrAlaAlaLeuLeuGlyLe 16
      ||||| ::|||::|||::|||::|||::|||::|||::|||
      649 GGTTGGTGCAGACATTGATTTCATCAGTCAGCGGCTGATTTTTT 698
      16 umetMetValValThrGly 22
      ||::|||::|||
      699 ATTCATGGTCTCTGTGGG 717

seq_name: /cgnl_7/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US01-15332A-39

seq_documentation_block:
```

```

; Sequence 39, Application PC/TUS0115332A
; GENERAL INFORMATION:
; APPLICANT: Brown, Joseph P.
; APPLICANT: Miller, Margaret
; APPLICANT: Burner, Glenna
; APPLICANT: Fabre-Suver, Christine
; APPLICANT: Pritchard, David
; APPLICANT: Lifespan Biosciences, Inc.
; TITLE OF INVENTION: Nucleic Acid Sequences for Novel GPCRs
; FILE REFERENCE: 017473-003710PC
; CURRENT APPLICATION NUMBER: PCT/US01/15332A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/203,217
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/205,945
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Ls191218 cluster
PCT-US01-15332A-39

alignment_scores:
    Quality: 45.00    Length: 23
    Ratio: 2.647      Gaps: 1
    Percent Similarity: 73.913    Percent Identity: 47.826

alignment_block:
US-09-215-435-225_COPY_1_22 x PCT-US01-15332A-39 ..
Align seg 1/1 to: PCT-US01-15332A-39 from: 1 to: 1176

2 GlyTrp.....ThrMetArgLeuValThrAlaAlaLeuLeuGlyLe 16
|||||.....|.....|.....|.....|.....|.....|
680 GGTGGTGTGACAGATGATGATTCATCAGCGGTGGTGGTATTTT 729

16 uMetMetValValThrGly 22
|.....|.....|.....|.....|.....|.....|
730 ATTCATGTTCTCTGTGGG 748

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-385-5

seq_documentation_block:
; Sequence 5, Application US/09978385
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2638
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(2520)

```

```

US-09-978-385-5

alignment_scores:
    Quality: 43.00    Length: 21
    Ratio: 2.688      Gaps: 0
    Percent Similarity: 76.190    Percent Identity: 38.095

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-978-385-5/rev ..
Align seg 1/1 to reverse of: US-09-978-385-5 from: 1 to: 2638

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||.....|.....|.....|.....|.....|.....|
2559 TTAGCATTTACATACAGGAAGATGACAGTCTTTGCTAAAGGAAGTCT 2510

17 tMetValValThr 21
|.....|.....|.....|.....|.....|.....|
2509 GAGCATCATCACT 2497

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-978-385-8

seq_documentation_block:
; Sequence 8, Application US/09978385
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2638
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(2520)
US-09-978-385-8

alignment_scores:
    Quality: 43.00    Length: 21
    Ratio: 2.688      Gaps: 0
    Percent Similarity: 76.190    Percent Identity: 38.095

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-978-385-8/rev ..
Align seg 1/1 to reverse of: US-09-978-385-8 from: 1 to: 2638

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||.....|.....|.....|.....|.....|.....|
2559 TTAGCATTTACATACAGGAAGATGACAGTCTTTGCTAAAGGAAGTCT 2510

17 tMetValValThr 21
|.....|.....|.....|.....|.....|.....|
2509 GAGCATCATCACT 2497

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-644-600-9

seq_documentation_block:

```

```

7 LeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetValValThrGly 22
:::||||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
188 GTGGTGCCTGGCAGCGCTGCTGATCGCCCTCCTCTTGGTCTTGGCTGGG 235

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-644-600-18

seq_documentation_block:
; Sequence 18, Application US/09644600
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; FILE REFERENCE: Overexpressed in Carcinomas
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 18
; LENGTH: 3147
; TYPE: RNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Antisense of TADG-15
US-09-644-600-18

alignment_scores:
Quality: 43.00 Length: 16
Ratio: 3.071 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 50.000

alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-644-600-18/rev ..

Align seg 1/1 to reverse of: US-09-644-600-18 from: 1 to: 3147

7 LeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetValValThrGly 22
:::||||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
2960 GTGGTGCCTGGCAGCGCTGCTGATCGCCCTCCTCTTGGTCTTGGCTGGG 2913

seq_name: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-965

seq_documentation_block:
; Sequence 965, Application US/09954456
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeut
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27

```

us-09-215-435-225_copy_1_22.rnpn

Thu Nov 15 10:52:13 2001

```
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 965
; LENGTH: 3175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-965
```

```
alignment_scores:
  Quality: 43.00      Length: 21
  Ratio: 3.071       Gaps: 0
  Percent Similarity: 66.667  Percent Identity: 38.095
```

```
alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-954-456-965/rev ..
Align seg 1/1 to reverse of: US-09-954-456-965 from: 1 to: 3175
```

```
2 GlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMe 18
|||||::: |||::: |||::: |||::: |||::: |||:::
2039 GCGTGGACTGTCTCATTGTCTCTGAGGCTGCTCCCTTTGAAATCTGTT 1990
```

```
18 tValValThrGly 22
:::|||| |||
1989 AATGGTGAGAGGC 1977
```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-954-456-214

```
seq_documentation_block:
; Sequence 214, Application US/09954456
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 214
; LENGTH: 3451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-214
```

```
alignment_scores:
  Quality: 43.00      Length: 21
  Ratio: 2.529       Gaps: 0
  Percent Similarity: 80.952  Percent Identity: 38.095
```

```
alignment_block:
US-09-215-435-225_COPY_1_22 x US-09-954-456-214 ..
Align seg 1/1 to: US-09-954-456-214 from: 1 to: 3451
1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||::: |||::: |||::: |||::: |||::: |||:::
1693 ATGGGCATCGTCATGTCCTCATCGCTCGCCATCGTGTGGCAATGT 1742
17 tMetValValThr 21
:::||||::: |||
1743 CCTGGTCATCACA 1755
```


THIS PAGE BLANK (USPTO)


```

17 tmetValValThrGly 22
|||||
141 GATGGTGGTCACTGGA 156

seq_name: gb_est93:BF853537

seq_documentation_block:
LOCUS BF853537 407 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-211200-002-h10 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853537
VERSION BF853537.1 GI:12241281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
211200-002-h10&t3=2000-12-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 407.
Location/Qualifiers
FEATURES
source
1..407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 87 a 104 c 124 g 92 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BF853537
Align seg 1/1 to: BF853537 from: 1 to: 407

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuGlyLeuMe 17
|||||
91 ATGGGTTGGACATGAGGCTGGTCACAGCAGCACTGTTACTGGGTCAT 140

17 tMetValValThrGly 22
|||||
141 GATGGTGGTCACTGGA 156

```

|||||
98 GATGGTGGTCACTGGA 113

seq_name: gb_est93:BF853928

seq_documentation_block:
LOCUS BF853928 536 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-261200-004-g12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853928
VERSION BF853928.1 GI:12241672
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 536)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
261200-004-g12&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 497.

FEATURES
source
1..536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 121 a 139 c 164 g 112 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BF853928 ..

Align seg 1/1 to: BF853928 from: 1 to: 536

1 MetGlyTrpThrMetArgLeuValThrAlaA1aLeuLeuGlyLeuMe 17
|||||
91 ATGGGGTTGGACATAGGCTGGTCACAGCAGCAGCTTTACTGGGTCTCAT 140
17 tMetValValThrGly 22
|||||

141 GATGGTGGTCACTGGA 156

seq_name: gb_est93:BF853000

seq_documentation_block:
LOCUS BF853000 558 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-f12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853000
VERSION BF853000.1 GI:12240848
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 558)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-f12&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 558.

FEATURES
source
1..558
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 124 a 146 c 163 g 124 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BF853000 ..

Align seg 1/1 to: BF853000 from: 1 to: 558

1 MetGlyTrpThrMetArgLeuValThrAlaA1aLeuLeuGlyLeuMe 17
|||||
93 ATGGGGTTGGACATAGGCTGGTCACAGCAGCAGCTTTACTGGGTCTCAT 142
17 tMetValValThrGly 22
|||||
143 GATGGTGGTCACTGGA 158

Thu Nov 15 10:52:13 2001

```

seq_name: gb_est93:BF852995
seq_documentation_block: 563 bp mRNA EST 16-JAN-2001
LOCUS BF852995
DEFINITION MR2-EN0093-191200-001-e01 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852995
VERSION BF852995.1 GI:12240843
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 563)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-e01&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 563.
Location/Qualifiers
1..563
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/notes="Organ: lung_normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 126 a 148 c 164 g 125 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BF852995
..
Align seg 1/1 to: BF852995 from: 1 to: 563

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
91 ATGGGTGGACATGAGGTGGTGCACAGCAGCACACTGTTACTGGGTCTCAT 140

17 tMetValValThrGly 22
|||||
141 GATGGTGGTCTCAGTGA 156

seq_name: gb_est93:BF852989
seq_documentation_block: 603 bp mRNA EST 16-JAN-2001
LOCUS BF852989
DEFINITION MR2-EN0093-191200-001-a12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852989
VERSION BF852989.1 GI:12240837
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 603)
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-a12&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 560.
Location/Qualifiers
1..603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/notes="Organ: lung_normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 140 a 157 c 171 g 135 t
ORIGIN

alignment_scores:
Quality: 109.00 Length: 22
Ratio: 4.955 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BF852989
..
Align seg 1/1 to: BF852989 from: 1 to: 603

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
91 ATGGGTGGACATGAGGTGGTGCACAGCAGCACACTGTTACTGGGTCTCAT 140

17 tMetValValThrGly 22
|||||
141 GATGGTGGTCTCAGTGA 156

seq_name: gb_est102:BG545668

```

```

seq_documentation_block:
LOCUS      BG545668          749 bp      mRNA          EST          04-APR-2001
DEFINITION 602572933f1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4700840 5',
            mRNA sequence.
ACCESSION  BG545668
VERSION    BG545668.1  GI:13544333
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 749)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1535 row: b column: 09
            High quality sequence stop: 740.
FEATURES   Location/Qualifiers
            source          1..749
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4700840"
                        /clone_lib="NIH_MGC_77"
                        /lab_host="DH10B (T1 phage-resistant)"
                        /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                        SfiI (ggccgctcgccc); Site_2: SfiI (ggccattatggcc); 5' and
                        3' adaptors were used in cloning as follows: 5' adaptor
                        sequence: 5'-CAGCGCCATTATGCC-3' and 3' adaptor sequence:
                        5'-ATTCTAGAGCGCGCGGCGGCGAGATG-dt(30)BN-3' (where B = A,
                        C, or G and N = A, C, G, or T). Average insert size 1.9
                        kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                        by PCR. This library was enriched for full-length clones
                        and was constructed by Clontech Laboratories (Palo Alto,
                        CA). Note: This is a NIH_MGC Library."
BASE COUNT  184 a 206 c 205 g 154 t
ORIGIN

alignment_scores:
Quality: 109.00      Length: 22
Ratio: 4.955        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_1_22 x BG545668 ..
Align seg 1/1 to: BG545668 from: 1 to: 749

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
86 ATGGGTGGACATGATGAGTGGTTACAGCAGCAGCTGTACTGGGTCTCAT 135

17 tMetValThrGly 22
|||||
136 GATGGTGGTCACCTGGA 151

seq_name: gb_est45:AW327171

seq_documentation_block:
LOCUS      AW327171          352 bp      mRNA          EST          10-JUL-2000
DEFINITION 20827 MARC 180V Bos taurus cDNA 5', mRNA sequence.
ACCESSION  AW327171
VERSION    AW327171.1  GI:6763092
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 454)

```

```

KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 352)
AUTHORS   Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M.,
            Bennett, G.A., Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Laegreid
            W.W. and Keefe, J.W.
TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
JOURNAL    Unpublished (2000)
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR Primers
            FORWARD: AGGAACACGCTATGACCAT
            BACKWARD: GTTTCACGTCACGACG
            Plate: 13 row: 0 column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES   Location/Qualifiers
            source          1..352
                        /organism="Bos taurus"
                        /db_xref="taxon:9913"
                        /clone_lib="MARC 180V"
                        /tissue_type="pooled"
                        /lab_host="DH10B"
                        /note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                        Library made from pooled tissue from lymph node, ovary,
                        fat, hypothalamus, and pituitary."
BASE COUNT  65 a 107 c 107 g 73 t
ORIGIN

alignment_scores:
Quality: 95.00      Length: 22
Ratio: 4.524        Gaps: 0
Percent Similarity: 95.455 Percent Identity: 86.364

alignment_block:
US-09-215-435-225_COPY_1_22 x AW327171 ..
Align seg 1/1 to: AW327171 from: 1 to: 352

1 MetGlyTrpThrMetArgLeuValThrAlaLeuLeuLeuGlyLeuMe 17
|||||
108 ATGGGTGGACCATGAGGTGGTGCACAGCAGCCCTGCTCTGGGCTCGC 157

17 tMetValThrGly 22
|||||
158 GGTGGCGGTCACCTGGA 173

seq_name: gb_est93:BF890431

seq_documentation_block:
LOCUS      BF890431          454 bp      mRNA          EST          18-JAN-2001
DEFINITION 292000 MARC 380V Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BF890431
VERSION    BF890431.1  GI:12281889
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 454)

```

AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrnkug,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.W. and Keele,J.W.

TITLE Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle

JOURNAL COMMENT Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTGACGACG
Plate: 57 row: E column: 23
Seq primer: ATTGTAGTGACTATATAG.

FEATURES

Location/Qualifiers	1..454
source	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/clone_lib="MARC 3BOV"
	/tissue_type="pooled"
	/lab_host="DH10B"
	/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
	Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 93 a 134 c 132 g 94 t 1 others

ORIGIN

```

alignment_scores:
Quality: 95.00 Length: 22
Ratio: 4.524 Gaps: 0
Percent Similarity: 95.455 Percent Identity: 86.364

alignment_block:
US-09-215-435-225_COPY1_1_22 x BF890431 ..
Align seg 1/1 to: BF890431 from: 1 to: 454

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
74 ATGGCTGGACCATGAGCGTGGTGCACAGCAGCCCTCTCTGGCGCTCGC 123

17 tMetValValThrGly 22
|||||
124 GGTCGGCTCACTGGA 139

seq_name: gb_est105:H83784

seq_documentation_block:
LOCUS H83784 440 bp mRNA EST 13-NOV-1995
DEFINITION ys64b02.r1 Soares retina N2b4HR Homo sapiens cDNA clone
IMAGE:219531 5', mRNA sequence..
ACCESSION H83784
VERSION H83784.1 GI:1062455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 440)
AUTHORS Hillier,L.M., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares M. Tap,F., Trevasik,E., Waterston ,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
```

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 364
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 911 Std Error: 0.00
Seq primer: M13Rp1
High quality sequence stop: 364.

FEATURES

Location/Qualifiers	1..440
source	/organism="Homo sapiens"
	/db_xref="GDB:3847524"
	/db_xref="taxon:9606"
	/clone="IMAGE:219531"
	/clone_lib="Soares retina N2b4HR"
	/sex="male"
	/tissue_type="retina"
	/dev_stage="55 year old"
	/lab_host="DH10B (ampicillin resistant)"
	/note="Organ: eye; Vector: pT7D3 (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCACTGAACTGGGAGCGGCCGCTTTTTTTTTTTTTTTT 3'],"
	double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7D3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)+ RNA was extracted hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 99 a 99 c 131 g 103 t 8 others

ORIGIN

```

alignment_scores:
Quality: 82.00 Length: 22
Ratio: 3.905 Gaps: 0
Percent Similarity: 95.455 Percent Identity: 95.455

alignment_block:
US-09-215-435-225_COPY1_1_22 x H83784 ..
Align seg 1/1 to: H83784 from: 1 to: 440

1 MetGlyTrpThrMetArgLeuValThrAlaAlaLeuLeuGlyLeuMe 17
|||||
44 ATGGCTGGACAATGAGCGTGGTCACACA.GCACTGTACTGGTCTCAT 92

17 tMetValValThrGly 22
|||||
93 GATGGCTGGCTCACTGGA 108

seq_name: gb_est93:BF846063

seq_documentation_block:
LOCUS BF846063 464 bp mRNA EST 16-JAN-2001
DEFINITION QVO-EN0055-181000-439-f11 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF846063
VERSION BF846063.1 GI:12233213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 464)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.
```


Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0055-181000-439-fil&t3=2000-10-18&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 23

High quality sequence stop: 464.

FEATURES

source

1. .464

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="EN0055"

/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

94 a 134 c 122 g 114 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 82.00 Length: 18

Ratio: 4.556 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_1_22 x BF846063/rev ..

Align seg 1/1 to reverse of: BF846063 from: 1 to: 464

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21

|||||

444 ATGAGGCTGGTGCACAGCAGCACTGTTACTGGGTCTCATGATGGTGTCAC 395

21 rGly 22

|||||

394 TGA 391

seq_name: gb_est93:BF846065

seq_documentation_block:

LOCUS BF846065 467 bp mRNA EST 16-JAN-2001

DEFINITION QV0-EN0055-181000-439-g06 EN0055 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF846065

VERSION BF846065.1 GI:12233215

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 467)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0055-181000-439-g06&t3=2000-10-18&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 467.

FEATURES

source

1. .467

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="EN0055"

/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

94 a 134 c 126 g 113 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 82.00 Length: 18

Ratio: 4.556 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_1_22 x BF846065/rev ..

Align seg 1/1 to reverse of: BF846065 from: 1 to: 467

5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21

|||||

447 ATGAGGCTGGTGCACAGCAGCACTGTTACTGGGTCTCATGATGGTGTCAC 398

21 rGly 22

|||||

397 TGA 394

seq_name: gb_est90:BF599823

seq_documentation_block:

LOCUS BF599823 413 bp mRNA EST 13-DEC-2000

DEFINITION 26147 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BF599823

VERSION BF599823.1 GI:11696691

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 413)

AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
,W.W. and Keele,J.W.
TITLE Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCAGTCACGAGC
Plate: 35 row: D column: 18
Seq primer: ATTAGGTGACACTATAG.
FEATURES
Location/Qualifiers
1..413
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3bov"
/tissue_type="pooled"
/lab_host="DH10b"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 87 a 126 c 107 g 93 t
ORIGIN

alignment_scores:
Quality: 68.00 Length: 18
Ratio: 4.000 Gaps: 0
Percent Similarity: 94.444 Percent Identity: 83.333
alignment_block:
US-09-215-435-225_COPY_1_22 x BF599823 ..
Align seg 1/1 to: BF599823 from: 1 to: 413
5 MetArgLeuValThrAlaAlaLeuLeuLeuGlyLeuMetMetValValTh 21
|||||
2 ATGAGGCTGGTCACAGCAGCCCTGCTCTGGGCTTCGGCGGTGGCGGTAC 51
21 rgly 22
||||
52 TGA 55

seq_name: gb_est93:BF848191

seq_documentation_block:
LOCUS BF848191 463 bp mRNA EST 16-JAN-2001
DEFINITION QV0-EN0055-031100-480-c10 EN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF848191
VERSION BF848191.1 GI:12235341
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 463)
Dias Neto,E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., Geoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

alignment_scores:
Quality: 64.50 Length: 18
Ratio: 3.794 Gaps: 1
Percent Similarity: 94.444 Percent Identity: 88.889
alignment_block:
US-09-215-435-225_COPY_1_22 x BF848191 ..
Align seg 1/1 to: BF848191 from: 1 to: 463
5 MetArgLeuValThrAlaAlaLeuLeuGlyLeuMetMetValValTh 21
|||||
30 ATGAGGCTGGTCACAGCAGCA...GTACTGGGTCTCATGATGGTGGTGCAC 76
21 rgly 22
||||
77 TGA 80

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0055-
031100-480-c10&t3=2000-11-03&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 463.
FEATURES
Location/Qualifiers
1..463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 119 c 136 g 92 t
ORIGIN

JOURNAL
MEDLINE
COMMENT

THIS PAGE BLANK (USPTO)

Thu Nov 15 10:52:13 2001

us-09-215-435-225_copy_23_227.ige

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
467 TTGAGGAGGAAAGTCACTCTCTCTCCCAAGGAAACAACTCGA 516
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
517 GGCTCTGGAAAATGACAGATTCTGAACCGCTTCCACCTGGGGCAAC 566
167 oGluAlaSerThrGlnPheMetThrGlnAsnTrpGlnAspSerProThrL 184
|||||
567 TGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGACTCACCACCC 616
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
617 TCCAGGCTCCAGAGAAAGGCCAGGAGCCCAAGCACAAAACCAAGGCG 666
201 GluLeuAlaLacys 205
|||||
667 GAGATAGCTGCCTGC 681

seq_name: gb_pat2:AX083422

seq_documentation_block: 897 bp DNA PAT 28-FEB-2001

LOCUS AX083422 Sequence 114 from Patent WO0112660.

DEFINITION AX083422

ACCESSION AX083422

VERSION AX083422.1 GI:13185262

KEYWORDS human.

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Kato,S. and Kimura,T.

TITLE Human proteins having hydrophobic domains and dnas encoding these

proteins

JOURNAL Patent: WO 0112660-A 114 22-FEB-2001;

SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)

FEATURES

source

1..897

/organism="Homo sapiens"

/db_xref="taxon:9606"

99..782

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC33305.1"

/db_xref="GI:13185263"

/translation="MGWTRLVTAALLGLMMVVTGDEENSPCAHEALLDDETLFCQ

GLEVFYFPELGNIGCKVVPDCNNYRQKITSWMEPIVKFPGAVDGATILVMVDPDAPSR

AEPRQFRHNLWTDIKGADLKKGIQGOELSAQAPSPHGFHYOFFYVLEQEGK

VISLLPKENKTRGSKMDRFLNRFHLGEPEASTQFMTQYQDSPTLQAPRERASEPKH

KNQAEIAC"

BASE COUNT 226 a 253 c 234 g 184 t

ORIGIN

alignment_scores:

Quality: 1115.00 Length: 205

Ratio: 5.439 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_23_227 x AX083422 ..

Align seg 1/1 to: AX083422 from: 1 to: 897

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17

|||||

165 GACGAGGATGAGAACACCGCTGTGCCCATGAGCCCTCTTGGACGAGGA 214

|||||

17 pThrLeuPheCysGlnGlyLeuGluValPheTrpProGluLeuGlyAsnI 34

|||||

215 CACCCTCTTTGCCAGGGCTTGAAGTTTCTACCCAGAGATTGGGGAACA 264

34 leGlyCysLysValValProAspCysAsnAsnTrpArgGlnLysIleThr 50
|||||
265 TTGGCTGCAAGGTGTTCTTGATTTAAACAACCTACAGACAGAGATCACC 314
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
315 TCCGTGGATGGAGCGATAGTCAAGTTCCCGGGGCGCTGGACGGCGCAAC 364
67 rTyriLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
365 CTATATCTCTGGTGTGGTGGATCCAGATGCCCTTAGCAGACGAGACCCA 414
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
415 GACAGAGATTCTGGAGACATTGGCTGTAAACAGATATCAAGGGCGCGAC 464
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTrpGlnAlaPr 117
|||||
465 CTGAAGAAAGGGAAGATTTCAGGGCCAGAGTTATCAGCCTACCAAGGCTCC 514
117 oSerProProAlaHisSerGlyPheHisArgTrpGlnPheValTrpL 134
|||||
515 CTCCCCACCGGACACAGTGGCTTCCATCGCTACCAAGTTCTTTGTCTATC 564
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
565 TTCAGAGAGGAAAGTCACTCTCTCTCTCCCAAGGAAACAAACTCGA 614
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
615 GGCTCTTGAAAATGACAGATTTCTGAACCGTTTCCACCTGGGGGAAC 664
167 oGluAlaSerThrGlnPheMetThrGlnAsnTrpGlnAspSerProThrL 184
|||||
665 TGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGGACTCACCACCC 714
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
715 TCCAGGCTCCAGAGAAAGGCCAGGAGCCCAAGCACAAAACCAAGGCG 764

201 GluLeuAlaAlaCys 205

|||||

765 GAGATAGCTGCCTGC 779

seq_name: gb_pat1:AX060293

seq_documentation_block:

LOCUS AX060293 932 bp DNA PAT 22-JAN-2001

DEFINITION Sequence 1 from Patent WO0078802.

ACCESSION AX060293

VERSION AX060293.1 GI:12405782

KEYWORDS human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 932)

Herrmann,J.L.

Shimkets,R.A., Fernandes,E., Vernet,C., Yang,M., Boldog,F.L. and

Patent: WO 0078802-A 1 28-DEC-2000;

JOURNAL Curagen Corporation (US)

FEATURES

Location/Qualifiers

1..932

/organism="Homo sapiens"

/db_xref="taxon:9606"

113..796

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC24866.1"

/db_xref="GI:12405783"

/translation="MGWTRLVTAALLGLMMVVTGDEENSPCAHEALLDDETLFCQ

GLEVFYBELNGICKVVPDCNNYRQKITSMWEPVIVKFGAVDGYATILVMVDPDPAPSR
AEPQRWRHLVTDIKGADLKEGKIQOELSAYQAPSPAHSFGHRYQFFVYLQEGK
VISLKPENKTRGSKWMDRLFNHLPEPEASTQFMTQYQDSPTLQAPRERASEPKH
KNOAEIAAC"
BASE COUNT 244 a 257 c 244 g 187 t
ORIGIN

alignment_scores:
Quality: 1111.00 Length: 205
Ratio: 5.420 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.512

alignment_block:
US-09-215-435-225_COPY_23_227 x AX060293 ..
Align seg 1/1 to: AX060293 from: 1 to: 932

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
179 GACGAGGATGAGAACAGCCGCTGTGCCCATGAGGCCCTCTTGACGAGGA 228
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
229 CACCCCTCTTTGCCAGGGCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 278
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
279 TTGGCTGCAAGGTGTCTCTGATTGTAACTATACAGACAGAAAGATCAC 328
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
329 TCCTGGATGAGCGGATAGTCAAGTTCCTCCGGGGCGCTGGACGGCGCAAC 378
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
379 CTATATCCTGGTATGGTGGATCCAGATGCCCTAGCAGAGCAGAACCCA 428
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
429 GACAGAGATTCGGAGACATTTGGCTGGTAACAGATATCAAGGCGCCGAC 478
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
479 CTGAGGAAGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACCAAGCTCC 528
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
529 CTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTTGTCTATC 578
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
579 TTCAGGAAGAAAGTCATCTCTCTCTCCCAAGGAAACAAACTCGA 628
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
629 GGCTCTTGGAAATGACAGATTTCTGAACCGTTTCCACCTGGCGCAACC 678
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
679 TGAAGCAAGCACCCAGTTTCATGACCCAGAACTACCAAGGACTCACCAACC 728
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
729 TCCAGGCTCCAGAGAAAGGGCCAGGAGCCCAAGCACAACCAACCAAGCG 778
201 GluIleAlaAlaCys 205
779 GAGATAGTGCCTGC 793

seq_name: gb_sts2:G27363
seq_documentation_block:

G27363 352 bp DNA STS 28-JUN-1996
DEFINITION human STS SHGC-31033, sequence tagged site.
ACCESSION G27363
VERSION G27363.1 GI:1396086
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Myers,R.M.
JOURNAL Unpublished (1996)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: TTGGAAGAGGAGGGTTCTG
Primer B: AAGCACAAAACCCAGGCG
STS size: 133
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H83927
-- Washington University/Merck EST sequence.

FEATURES
source

1..352
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"

STS
primer_bind 4..136
primer_bind 4..23

BASE COUNT 69 a 76 c 110 g 96 t 1 others
ORIGIN

alignment_scores:

Quality: 444.00 Length: 83
Ratio: 5.349 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_23_227 x G27363/rev ..

Align seg 1/1 to reverse of: G27363 from: 1 to: 352

123 SerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGluGlyLysVa 139
|||||
352 AGTGGCTTCATCGCTACCAAGTCTTTGTCTATCTTCAGGAGGAAAAAGT 303

```

139 lileserLeuLeuProLysGluAsnLysThrArgGlySerTrpLysMeta 156
|||||
302 CATCTCTCTCTCCCAAGCAAAACAACTCGAGCTCTTGCAAAATCG 253
|||||
156 spArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSerThrGln 172
|||||
252 ACAGATTCTGAACCGTTTCACCTGGCGGACMCCTGAAGCAAGACCCAG 203
|||||
173 PheMetThrGlnAsnTyrGlnAspSerProThrLeuGlnAlaAlaProArgG1 189
|||||
202 TTCATGACCAGACACTACAGGACTCACCACCTCCAGGCTCCAGAGA 153
|||||
189 uArgAlaSerGluProLysHisLysAsnGlnAlaGluIleAlaAlaCys 205
|||||
152 AAGGGCAGCGAGCCCAAGCAACAAAACAGGCGGAGATAGCTGCCTGC 104

seq_name: gb_htg13:AC055854

seq_documentation_block:
LOCUS AC055854 179814 bp DNA HTG 25-MAY-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-459E5 map 8, WORKING DRAFT
SEQUENCE, 26 unordered pieces.
ACCESSION AC055854
VERSION AC055854.3 GI:8084680
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 179814)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-459E5
2 (bases 1 to 179814)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferrelira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Minova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,P., O'Donnell,P.,
O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrelli,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye.W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7770648.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: I9765
Center clone name: 459_E_5
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

FEATURES
Source
1..179814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1730: contig of 1730 bp in length
* 1731 1830: gap of 100 bp
* 1831 3230: contig of 1400 bp in length
* 3231 3330: gap of 100 bp
* 3331 5675: contig of 2345 bp in length
* 5676 5775: gap of 100 bp
* 5776 7874: contig of 2099 bp in length
* 7875 7974: gap of 100 bp
* 7975 10477: contig of 2503 bp in length
* 10478 10577: gap of 100 bp
* 10578 12553: contig of 1976 bp in length
* 12554 12653: gap of 100 bp
* 12654 16386: contig of 3733 bp in length
* 16387 16486: gap of 100 bp
* 16487 19715: contig of 3229 bp in length
* 19716 19815: gap of 100 bp
* 19816 23166: contig of 3351 bp in length
* 23167 23266: gap of 100 bp
* 23267 28853: contig of 3589 bp in length
* 28856 28953: gap of 100 bp
* 28956 31690: contig of 4735 bp in length
* 31691 31790: gap of 100 bp
* 31791 38834: contig of 7044 bp in length
* 38835 38934: gap of 100 bp
* 38935 46670: contig of 7736 bp in length
* 46671 46770: gap of 100 bp
* 46771 52635: contig of 5865 bp in length
* 52636 52735: gap of 100 bp
* 52736 59853: contig of 7118 bp in length
* 59854 59953: gap of 100 bp
* 59954 66887: contig of 6934 bp in length
* 66888 66987: gap of 100 bp
* 66988 75106: contig of 8119 bp in length
* 75107 75206: gap of 100 bp
* 75207 84076: contig of 8870 bp in length
* 84177 91487: contig of 7311 bp in length
* 91488 91587: gap of 100 bp
* 91588 98734: contig of 7147 bp in length
* 98735 98834: gap of 100 bp
* 98835 107787: contig of 8953 bp in length
* 107788 107887: gap of 100 bp
* 107888 118191: contig of 10304 bp in length
* 118192 118291: gap of 100 bp
* 118292 134239: contig of 15948 bp in length
* 134240 134339: gap of 100 bp
* 134340 148401: contig of 14062 bp in length
* 148402 148501: gap of 100 bp
* 148502 163397: contig of 14896 bp in length
* 163398 163497: gap of 100 bp
* 163498 179814: contig of 16317 bp in length.
Location/Qualifiers
1..179814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"

```



```

CDS
/gene="PBP"
/29..592
/gene="PBP"
/codon_start=1
/product="phosphatidylethanolamine-binding protein"
/protein_id="CAA50708.1"
/db_xref="GI:510339"
/db_xref="SWISS-PROT:P31044"
/translation="MAADISQWAGPLSLQEVDEPPQHALRYDYGVGVTVDELGKVLPT
QVMRPSISWDGLDPEKLYTLVLTDPADPSRKDKFRWEHHFLVYNNKGNDISSGTV
LSEVYGGSPKPTGLHRYVWLVEYQDPLNCDPEILSNKSGDNRGKFKVESFRKKYHL
GAPVAGTCFOAEWDDSVPKLHDQLAGK"
3'UTR
590..1037
variation
766
/note="in 6 clones"
/replace="t"
BASE COUNT      244 a      299 c      290 g      204 t
ORIGIN

alignment_scores:
  Quality: 237.50      Length: 185
  Ratio: 2.284      Gaps: 6
Percent Similarity: 56.216      Percent Identity: 32.973

alignment_block:
US-09-215-435-225_COPY_23_227 x RNPBP ..

Align seg 1/1 to: RNPBP from: 1 to: 1037

30  GluLeuGlyAsnIleGlyCysLysValProAspCysAsnAspTyrAr 46
|||||
134  GAGCTGGGCAAGTG.....CTGACGCCACCACCGAGTGCATGAATAG 174

46  gGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAlaVal 63
|||
175  ACCAAGCAGCATTTTCATGG.....GATGCCGTGTG 203

63  aIAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 79
|||||
204  ATCCTGGGAAGCTCTACACCCTGGTCTCACAGACCCCGATGCTCCAGC 253

80  ArgAlaGluProArgGlnArgPheThrArgHisTrpLeuValThrAspIle 96
|||
254  AGGAAGGACCCCAATTCCAGGAGTGGCACCACCTCTCTGGTGGTCAACAT 303

96  eLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuSerA 113
|||||
304  GAAGGCAACGACATTAGCAGTGCACTGTC.....CTCTCCG 341

113  latYrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGln 129
|||
342  AATACGTGGGCTCCGGACCTCCCAAAGACACAGGTCTGCACCGCTACGTC 391

130  phePheValTyrLeuGlnGluGlyLysValIleSer.....Ie 142
|||||
392  TGGCTGGTGTATGACGACGAGCAGCCCTCTGAACGTGACGAGCCCATCTC 441

142  uLeuProLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheL 159
|||||
442  CAGCACACAGTCTGAGACACACCGCGGCAAGTTCAGGTGGAGTCTCTCC 491

159  euAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThr 175
|||||
492  GCAAGAAGTACCACTTGGGAGCCCGGTGGCCGACGCTGCTTCCAGGCA 541

176  GlnAsnTyrGlnAspSer..... 181
|||
542  GAG...TGGGATGACTCTCTGTGCCAAGCTGCACGATCAGCTGGGCTGGGA 588

182  .....ProThrLeuGlnAlaP 197
|||||
589  GTAGGGGCGCTGCAGAGCCCGCAGCCCGGGGACCCACACAGTACAGTCAA 638

```



```

176  GlnAsnTyrGlnAspSer..... 181
    :: :::::|||||
567  GAG...TGGGATGACTCTGTGCCCAAGCTGCATCAGCTGGCTGGGAA 613
    |||||:::|:::
182  .....ProThrLeuGlnAlap 187
    |||||:::|:::
614  GTAGGGCGCTGCAGCGCCGAGCCCGGGGAGCCCACTACAGTCAA 663
    |||||:::|:::
187  roArg 188
    |||
664  CTCGT 668

```

```

name: gb_htg12:AC037441
Documentation_block:
5 AC037441 162795 bp DNA HTG 03-MAR-2001
UNION Homo sapiens clone RP11-87E22, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
MISSION AC037441
ION AC037441.3 GI:13184260
ORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
EE human.
GANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 162795)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens Chromosome, clone RP11-87E22

```

Campaniano, A., Castle, A., Choepiel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S.B., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kardas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Llorente, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuck, A., McKernan, K., MCPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivier, T.M., Olivet, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Tridillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (08-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:8705171.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://www.genome.washington.edu/rw/RepeatMasker.htm>

```

Center project name: L9203
Center clone name: 87_E_22
----- Summary Statistics -----
Sequencing vector: M13; M77815; 81% of reads
Sequencing vector: plasmid; n/a; 19% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157025 bases at least 040

```

```

Ratio: 5.488                      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 97.674

alignment_block:
US-09-215-435-225_COPY_23_227 x AC037441/rev ..

Align seg 1/1 to reverse of: AC037441 from: 1 to: 162795

22 GInGlyLeuGluValPheTyrProGluLeuGlyAsnIleGlyCysLysVa 38
:::|||||
151900 AGGGCCCTGAAGTTTCTACCCAGAGTTGGGACATGGCTGCAAGT 151851

38 lValProAspCysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluP 55
|||||
151850 TGTTCCTGATTGAACTAACAACACAGACAGAGATCACCTCCTGGATGAGC 151801

55 roIleValLysPheProGlyAlaValAsp 64
|||||
151800 CGATAGTCAAGTTCCCGGGCGCGTGAC 151772

seq_name: gb_htgl7:AC084081

seq_documentation_block:
LOCUS AC084081 169645 bp DNA HTG 11-MAR-2001
DEFINITION Homo sapiens chromosome 8 clone RP11-89M8 map 8, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
ACCESSION AC084081
VERSION AC084081.3 GI:13273380
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 169645)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boquslavskiy,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,
Lanazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPherson,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo.A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (12-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 11, 2001 this sequence version replaced gi:12583824.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11307
Center clone name: 89_M_8

```

```

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162624 bases at least Q40
Consensus quality: 165742 bases at least Q30
Consensus quality: 167046 bases at least Q20
Insert size: 169000; agarose-fp
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 335: contig of 335 bp in length
* 336 435: gap of 100 bp
* 436 2289: contig of 1854 bp in length
* 2290 2389: gap of 100 bp
* 2390 4997: contig of 2608 bp in length
* 4998 5097: gap of 100 bp
* 5098 64677: contig of 59580 bp in length
* 64678 64777: gap of 100 bp
* 64778 66151: contig of 3374 bp in length
* 68152 68251: gap of 100 bp
* 68252 72655: contig of 4404 bp in length
* 72656 72755: gap of 100 bp
* 72756 78893: contig of 6138 bp in length
* 78894 78993: gap of 100 bp
* 78994 84664: contig of 5671 bp in length
* 84665 84764: gap of 100 bp
* 84765 90853: contig of 6089 bp in length
* 90854 90953: gap of 100 bp
* 90954 98425: contig of 7472 bp in length
* 98426 98525: gap of 100 bp
* 98526 106371: contig of 7846 bp in length
* 106372 106471: gap of 100 bp
* 106472 115673: contig of 9202 bp in length
* 115674 115773: gap of 100 bp
* 115774 127138: contig of 11365 bp in length
* 127139 127238: gap of 100 bp
* 127239 138436: contig of 11198 bp in length
* 138437 138536: gap of 100 bp
* 138537 149766: contig of 11230 bp in length
* 149767 149866: gap of 100 bp
* 149867 167340: contig of 17474 bp in length
* 167341 167440: gap of 100 bp
* 167441 169645: contig of 2205 bp in length.
FEATURES
source
Location/Qualifiers
1..169645
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"
/chromosome="8"
/misc_feature
1..335
/clone="RP11-89M8"
/clone_lib="RPC1-11 Human Male BAC"
1..335
/note="assembly_fragment
clone_end:SP6
vector_side:left"
436..2289
/note="assembly_fragment"
2390..4997
/note="assembly_fragment"
5098..64677
/note="assembly_fragment"
64778..68151
/note="assembly_fragment"

```

Thu Nov 15 10:52:13 2001

```

misc_feature      6852..72655
                  /note="assembly_fragment"
misc_feature      72756..78893
                  /note="assembly_fragment"
misc_feature      78994..84664
                  /note="assembly_fragment"
misc_feature      84765..90853
                  /note="assembly_fragment"
misc_feature      90954..98425
                  /note="assembly_fragment"
misc_feature      98526..106371
                  /note="assembly_fragment"
misc_feature      106472..115673
                  /note="assembly_fragment"
misc_feature      115774..127138
                  /note="assembly_fragment"
misc_feature      127239..138436
                  /note="assembly_fragment"
misc_feature      138537..149766
                  /note="assembly_fragment"
misc_feature      149867..167340
                  /note="assembly_fragment"
misc_feature      167441..169645
                  /note="assembly_fragment
                  vector_side:right"
BASE COUNT      44253 a 41074 c 39720 g 42960 t 1638 others
ORIGIN

alignment_scores:
  Quality: 236.00      Length: 43
  Ratio: 5.488         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.674

alignment_block:
US-09-215-435-225_COPY_23_227 x AC084081/rev ..
Align seg 1/1 to reverse of: AC084081 from: 1 to: 169645
22 GlnGlyLeuGluValPheTrpProGluLeuGlyAsnIleGlyCysLysVa 38
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41905 AGGGCCCTTGAAGTTTCTACCCAGAGTGGGGACATGGCTGCAAGGT 41856

38 lValProAspCysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluP 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
41855 TGTTCTGATTGTAACACTACAGACAGACAGATCACCTCTGGATGGAGC 41806

55 rolleValLysPheProGlyAlaValasp 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
41805 CGATAGTCAAGTTCGGGGGGCGGTGGAC 41777

seq_name: gb_in4:OVD1PROTN

seq_documentation_block:
LOCUS      OVD1PROTN      611 bp      mRNA
DEFINITION O.volvulus mRNA for dl protein.
ACCESSION  X87991
VERSION    X87991.1 GI:1143526
KEYWORDS   D1 gene; D1 protein.
SOURCE     Onchocerca volvulus.
ORGANISM   Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE  1 (bases 1 to 611)
AUTHORS   Erttmann,K.D. and Gallin,M.Y.
TITLE     Onchocerca volvulus: identification of cDNAs encoding a putative
            phosphatidyl-ethanolamine-binding protein and a putative partially
            processed mRNA precursor
JOURNAL   Gene 174 (2), 203-207 (1996)
MEDLINE   97045813
REFERENCE  2 (bases 1 to 611)
AUTHORS   Erttmann,K.D.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (16-JUN-1995) K.D. Erttmann, Bernhard Nocht Inst for
            Tropical Med., Bernhard-Nocht-Str. 74, 20359 Hamburg, FRG
COMMENT    Related sequences M27807, U05980, U05981, U05982, U05984, U05985,
            U05243, U05244.
FEATURES   Location/Qualifiers
            source          1..611
                        /organism="Onchocerca volvulus"
                        /db_xref="taxon:6282"
                        /dev_stage="adult"
                        /clone_lib="lambda ZAP"
                        /cell_line="E.Coli"
            gene            1..459
                        /gene="dl"
            CDS              <1..459
                        /gene="dl"
                        /codon_start=1
                        /product="D1 protein"
                        /protein_id="CAA61244.1"
                        /db_xref="GI:1143527"
                        /db_xref="SWISS-PROT:P54186"
                        /translation="VYVSTSPTKLVNVSNNLTNVLNGLTPTQVKNOPTKYSWDAEPG
                        ALYTLVMTDPDAPSRKNPVRERHWHLLINISQNVSSGTVLSVDYIGSGPPKGTGLHR
                        YVFLVYKQPGSITDTHQGNRPNFVKVMDFAHKLGNPVAGNFFQAKHED"
BASE COUNT      196 a 114 c 123 g 178 t
ORIGIN

alignment_scores:
  Quality: 234.50      Length: 152
  Ratio: 2.393         Gaps: 6
Percent Similarity: 64.474 Percent Identity: 36.842

alignment_block:
US-09-215-435-225_COPY_23_227 x OVD1PROTN ..
Align seg 1/1 to: OVD1PROTN from: 1 to: 611
33 AsnIleGlyCysLysValValPro...AspCysAsnAsnTyrArgGlnLy 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58 ANTCTGGGCAATGAATCTACCGCGAGGTAAAGATCAGCCGACAAA 107

48 sileThrSerTrpMetGluProIleValLysPheProGlyAlaValaspG 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
108 AGTA...TCATGG.....GATCGGGAACCTG 130

65 lValAlaThrTyrIleLeuValMetValaspProAspAlaProSerArgAla 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
131 GAGCCTTATATACGCTCGTTATCACTGATCGGAGCACCACCTCGAAAA 180

82 GluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGl 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
181 AACCCCGTATTACAGAGTGGCACCATTGGTGAATAATAATATTCTCG 230

98 yAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 ACAAAATGTAGCAGTGGCACAGTG.....TTATCTGATTATA 268

115 lNAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePhe 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
269 TTGGATCAGTCCACCACCAAAAGGACAGGACTTCATCGTTATGTTCTTG 318

132 ValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLy 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
319 GTTTATAAACAACTGGAGTATCACCGAT.....ACTCAACATGGCGG 362

148 sThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuG 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
363 AAATCGCCCAATTTCAAGATTATGATTGTTTTCACCAACAAACATCATTG 412

165 lYGlupProGluAlaSerThrClnPheMetThrGlnAsn.....Tyr 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
413 GAAATCCAGTTGCCGGGAACCTTCTCCAGGCTAAACATGAAGACTAATAT 462

```

179 GlnAsp 180
:::||||
463 GAAGAC 468

seq_name: gb_in3:CELF40A3

seq_documentation_block:

LOCUS CELF40A3 41155 bp DNA INV 02-AUG-1997
DEFINITION Caenorhabditis elegans cosmid F40A3.
ACCESSION AF016423
VERSION AF016423.1 GI:2291195

KEYWORDS

SOURCE Caenorhabditis elegans strain=Bristol N2.

ORGANISM

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 41155)

REFERENCE
AUTHORS Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Latreille,P., Lighning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Shownkeen,R., Smalton,N., Smith,A., Sonhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohldman,P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans

TITLE Nature 368 (6466), 32-38 (1994)

JOURNAL
MEDLINE
94150718

REFERENCE 2 (bases 1 to 41155)

AUTHORS Geisel,C., Bradshaw,H. and Keppler,D.

TITLE The sequence of C. elegans cosmid F40A3

JOURNAL Unpublished (1997)

REFERENCE 3 (bases 1 to 41155)

AUTHORS Waterston,R.

TITLE Direct Submission

JOURNAL Submitted (29-JUL-1997)

COMMENT Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1HQ, England
e-mail: rw@nematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is C37C3, 200 bp overlap; 3' lies in a gap. Actual start of this cosmid is at base position 197 of CELF40A3; actual end is at 41155 of CELF40A3

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

FEATURES source

Location/Qualifiers

1. 41155
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="v"
/clone="F40A3"
/complement(156..1605)
/gene="F40A3.5"
/complement(join(156..216,268..662,714..787,864..981,1094..1281,1357..1431,1476..1605))
/note="F40A3.5"
/gene="contains similarity to a protein kinase domain"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65324.1"
/db_xref="GI:2291201"
/translation="MELDHTPPPSVLNDCNSASYMTPYATVIAMSGLYLLAIFYFCKKSKMCPMSDSIYPQKRLTQRELKFDVRAHSVFAELKNYLIDESIEVDQFQIGQTADGIFRGVFPKTRNENAKVTAVKISFIVSKSISLELALRSLKLDHPNLIRLLAYSLSFTVFRPMIALEWPGGTADYFOFKVREKDDSPSPQLKMLSLYQVSLQALYIHSQLDDEGQELTHGRIFTRNVLTPEDLRKCEVKLGDFGDMGLGYSTPIIAYMPEILCCAEIRPPHPENDVMVGVFIWECLTLGAQPHFRKSVVEIKKSFRLPDRGLSCPPPTCPLDV"
6740..7594
/gene="F40A3.4"
/complement(6740..6827,7331..7394,7450..7594)
/gene="F40A3.4"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65323.1"
/db_xref="GI:2291200"
/translation="MSRMDSYSSSIREQKAILNGLLGLFADGIIYDNYDTTRKIHTVDGIYRMVALLIIVALLIIIAVFLICFFHCRANNEERSPLHSNRKSHFIEI"
8506..9283
/gene="F40A3.3"
/complement(8506..8600,8667..8781,8828..9283)
/gene="F40A3.3"
/note="Similar to phosphatidylethanolamine binding protein; coded for by C. elegans cDNA yk106a12.5; coded for by C. elegans cDNA yk106a12.3"
/codon_start=1
/protein_id="AAB65322.1"
/db_xref="GI:2291199"
/translation="MVVLTSLPALFFASRAPFAAATTSARFORGLATWAAEAFTHKHEVTPDLASNPSPKSVYKNSVEANLGNVLTPTQVKDTPKWDAPKAEALYTLIKTDPAPSRKEPTIREWHWLVNIPGNDIAGDITLSEYIGAPPPKTLGHRIVYLIYKQSGRIEDAEHGRLTNTSGDKRGWKAADFVAKHKLGAPEVFGNLFOAEYDDYVPLNKGALGA"
9824..12167
/gene="F40A3.2"
/complement(9824..9917,10735..11058,11837..12034,12091..12167)
/gene="F40A3.2"
/note="contains similarity to EGF-like domains"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65320.1"
/db_xref="GI:2291197"
/translation="MLRKLISAILIVLVACVLRKKKPCINGTPEGDRCYCIEGWGTGLCHRMVYSGFERLQNGSCIECELGWAGTDCDIIDCHGNPNYDLTCTCTVPSGKYCEIADTKDIYKWNNTSSIGPIGILTIIPLATIMCEHFARQLRLRVEOHLNGYITTRKQVDKMLVGLLEASDEKSVKSVKIAADSGDKEDKKEEEVVKRNSATTS EKKALLADPADM"
complement(23146..24752)
/gene="F40A3.6"
/complement(join(23146..23311,23512..23597,23644..23841,23885..24052,24594..24752))
/gene="F40A3.6"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB65325.1"
/db_xref="GI:2291202"
/translation="MSEYLGWNRKPAAGSRGNSFFORDIVPGGAHTKSYSEYKNIFF

```

127 ArcTyrGlnPhePheValTyrLeuGluGlnGlutLysValIleSer..... 141  
||||| :::: ||||| |::: ||||| :  
9074 CGTTACGTTCCTCATCAGCAACTCGCACGCATCCGAAGACGCCGA 9123  
  
142 .....LeuLeuProLysGluAsnLysThrArgGlySerTrpLysM 155  
||| :::: |::::::::::|  
9124 GCACGGACGCTTACCACAACACTTCGGGAGACAAAGAAGAGGATTGAAGG 9173  
  
155 etASPArGpheLeuASNArGPHeHisLeuGlyGLUProGluAlaSerThr 171  
||| :::: |::: ||| :  
9174 CTGCCGATTTTCGTCGCAAGACACAACTCGTGCTCCAGTCTTCGGAAT 9223  
  
172 GlnPheMetThrGlnAsnTyrrGlnAasp 180  
||| :::: ||:::||  
9224 CTGTTCACGCGCGAG...TAGCAGCAG 9247
```

```
seq_name: gb_htgl.AC006915

seq_documentation_block:
LOCUS      AC006915    134095 bp          DNA              HTG       24-F-
DEFINITION Caenorhabditis elegans clone Y97E10y, *** SEQUENCING I
***, 1 unordered pieces.
ACCESSION   AC006915
VERSION     AC006915.1 GI:4263433
KEYWORDS    HTG; HTGS_PHASE1
SOURCE      Caenorhabditis elegans.
ORGANISM    Caenorhabditis elegans.
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis
REFERENCE   1 (bases 1 to 134095)
AUTHORS     Waterston,R.H.
TITLE        The sequence of Caeohorhabditis elegans clone
JOURNAL      Unpublished
ENTRY       2 (bases 1 to 134095)
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (24-FEB-1999) Genome Sequencing Center, Wash
            University School of Medicine, 4444 Forest Park Parkway
            MO 63108, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currentl
            * consists of 1 contigs. The true order of the pieces
            * is not known and their order in this sequence record
            * arbitrary. Gaps between the contigs are represented
            * runs of N, but the exact sizes of the gaps are unkno
            * This record will be updated with the finished sequen
            * as soon as it is available and the accession number
            * be preserved.
FEATURES             Location/Qualifiers
     source           1..134095
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
                     /clone="Y97E10y"
BASE COUNT         43230 a 24499 c 23702 g 42664 t
ORIGIN
alignment_scores:
Quality:           230.00 Length:           159
Ratio:             2.447 Gaps:                7
Percent Similarity: 59.119 Percent Identity: 37.107

alignment_block:
US-09-215-435-225_COPY_23_227 x AC006915 ..
Align seg 1/1 to: AC006915 from: 1 to: 134095

27 PheTYrPrOGluLeuGlyASnllEGlyCYSLYSVALPRoAScYSAas 43
||| :::: ||||| |::: ||||| :
99439 TITCAGSGTAATCTCGAAAACGTT.....CTCACCTCCAACC.... 99474
```



```

alignment_block:
US-09-215-435-225_COPY_23_227 x AC012978/review ..
Align seg 1/1 to reverse of: AC012978 from: 1 to: 63551

54 GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 70
||||| ||||| :||| :||| :||| :||| :||| :||| :|||
23000 GAGCGCTGGTGAAATGGGAG...GCCGACGCCAATAGCTGTACACCT 22954
||||| :||| :||| :||| :||| :||| :||| :||| :|||

70 uValMetValAspProAspAlaProSerArgAlaGluProArgGlnAtp 87
| ||| :||| :||| :||| :||| :||| :||| :||| :|||
22953 CTGCATGACCATTCCGATGCCCATCTCGCAAGATCCCAGTTTAGGG 22904
||||| :||| :||| :||| :||| :||| :||| :||| :|||

87 heTrpArgHisTriPLeuValThrAspIleLysGlyAlaAspLeuLys 103
||||| :||| :||| :||| :||| :||| :||| :||| :|||
22903 AGTGCACCATTTGGCTGGTGGCAACAATACCCGCTGGAGATGTCGCCAAG 22854
||||| :||| :||| :||| :||| :||| :||| :||| :|||

104 GlyLysIleGInGlyGlnLeuSerAlaTyrGlnAlaProSerProPrp 120
||||| :||| :||| :||| :||| :||| :||| :||| :|||
22853 GCGGAGGTT.....CTCCCGCCTACGTGGGATCCGGCCCTCC 22816
||||| :||| :||| :||| :||| :||| :||| :||| :|||

120 cAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGlug 137
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
22815 ACCAGACACCGGACTCCATCGTTAGCTTTCTCTGATCTACGAGCACGGT 22766
||||| :||| :||| :||| :||| :||| :||| :||| :|||

137 lyLysValIle.....SerLeuLeuProLySgluAsn...LysThr 149
||||| :||| :||| :||| :||| :||| :||| :||| :|||
22765 GCAAGCTCAcATTCACGAGAGAAGCGACTGCCAATAACAGCGGAGATGA 22716
||||| :||| :||| :||| :||| :||| :||| :||| :|||

150 ArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyCl 166
||||| :||| :||| :||| :||| :||| :||| :||| :|||
22715 CCGGTGGCTTCAAATCCCGAGTTCGCCAAGAAGTACGCCCTCGGCA 22666
||||| :||| :||| :||| :||| :||| :||| :||| :|||

166 uProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAsp 180
||||| :||| :||| :||| :||| :||| :||| :||| :|||
22665 TCCCATTCGGGTAACTCTACAGGCTGAG...TACGACGAC 22626

seq_name: gb_inl.AC005891

seq_documentation_block:
LOCUS AC005891 136933 bp DNA INV 30-
DEFINITION Drosophila melanogaster, chromosome 2L, region 33F3-3
DS08479 and DS07153, complete sequence.
ACCESSION AC005891 AC005108 AC005109 AC005110
VERSION HTG.
KEYWORDS AC005891.1 GI:3810575
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 136933)
AUTHORS Ceilinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harri-
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Pa-
Preifer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H.,
Swirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhan-
Zleran,L.L. and Rubin,G.M.
Sequencing of Drosophila chromosome 2L, region 33F3-3
Unpublished (1998)
REFERENCE 2 (bases 1 to 136933)
AUTHORS Ceilinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harri-
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Pa-
Preifer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H.,
Swirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhan-

```

```

FEATURES
  source
    Location/Qualifiers
      i. 136953
        /organism="prosophila melanogaster"
        /strain="y2; cn bw sp"
        /db_xref="taxon:7227"
        /chromosome="2L"
        /map="33F3-34A2"
        /clone_lib="P1 library, partial Sau3A in pAd10sacBII"
        /clones="P1S D508479 (D303) and D507153 (D305)"
        /note="D508479 (D303) and D507153 (D305) were completed as
a project. D508479 extends from p1 end at bp 1 to p1 end
at bp 83797. D507153 extends from p1 end at bp 51,253 to
p1 end at bp 136,933."
      37380 a 28396 c 29033 g 42124 t
BASE COUNT
ORIGIN

```

```

alignment_scores:      Length: 131
      Quality: 229.00      Gaps: 5
      Ratio: 2.516
      Percent Similarity: 69.466      Percent Identity: 40.458

alignment_block:
US-09-215-435-225_COPY_23_227 x AC005891  ..

Align seg 1/1 to: AC005891 from: 1 to: 136933

54  GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 70
||||| ||||||| ||| ::| ||| |||
24363 GAGCCCTGGCTGAAATGGGAG...GCGAGCGCAATAAGCTGTACACCCT 24409
||||| ||||||| ||| ::| ||| |||

70  uValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgP 87
||||| ||||||| ||| ::| ||| |||
24410 CTGCATGACCGATCCGATCGGCCAGTGCAGAGGATCCCAAGTTTAGGG 24459
||||| ||||||| ||| ::| ||| |||

87  heTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 103
||||| ||||||| ||| ::| ||| |||
24460 AGTGGCACCAATTGCTGGGGCAACATCCCGTGGAGATGTCGGCCAA 24509
||||| ||||||| ||| ::| ||| |||

104  GlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProPr 120
||||| ||||||| ||| ::| ||| |||
24510 GCGCAGGTT.....CTCTCGGCCTACGTGGGATCCGGGCTCC 24547
||||| ||||||| ||| ::| ||| |||

120  oAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnGluG 137
||||| ||||||| ||| ::| ||| |||
24548 ACCAGACACCGGACTCCATCGTTACGTTTTCCTGATCTCAGCAGCGGT 24597
||||| ||||||| ||| ::| ||| |||

137  lYlsValIle.....SerLeuLeuProLysGluAsn...LysThr 149
||||| ||||||| ||| ::| ||| |||
24598 GCAAGCTCACATTCGACGAGAGCGACTGTGCCAATAACAGCGGAGATGGA 24647
||||| ||||||| ||| ::| ||| |||

150  ArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyG 166
||||| ||||||| ||| ::| ||| |||
24648 GCGCGTGGCTTCAAAATCGCGAGTTCGCCAAGAAGTACGCCCTTCGGCAA 24697
||||| ||||||| ||| ::| ||| |||

166  uProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAsp 180
||||| ||||||| ||| ::| ||| |||
24698 TCCCATTCGCGGTAACCTCTACCGAGGCTGAG...TAGCAGCGAC 24737
||||| ||||||| ||| ::| ||| |||

```

OM of: US-09-215-435-225_COPY_23_227 to: N_Geneseq_0601.* out_format : pfs

Date: Nov 15, 2001 4:29 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2157/app_query.fasta_1.1519
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=2000000000
-USBR=US09215435_eCGNL_1_447 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_23_227

Query length: 205

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 1043.840000

score_list:

Sequence	Strid	Orig	zScore	EScore	Len	! Documentation
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480				1115.00	2318.95	3.4e-121
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				1115.00	2316.86	4.4e-121
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1115.00	2316.86	4.4e-121
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1115.00	2316.86	4.4e-121
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480				1115.00	2315.97	5.0e-121
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AX37660				1115.00	2307.16	1.5e-120
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1109.00	2301.90	3.0e-120
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1106.00	2296.94	5.7e-120
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1104.00	2292.81	9.7e-120
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1089.00	2260.76	5.9e-118
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				1080.00	2242.61	6.1e-117
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				543.00	1123.27	1.4e-54
/cgnl_9/gcgdata/geneseq/geneseq/NA1992.DAT:AX37660				237.50	472.99	2.2e-18
/cgnl_9/gcgdata/geneseq/geneseq/NA1994.DAT:AX37660				237.50	472.99	2.2e-18
/cgnl_9/gcgdata/geneseq/geneseq/NA1995.DAT:AX37660				205.00	405.28	1.3e-14
/cgnl_9/gcgdata/geneseq/geneseq/NA1992.DAT:AX37660				205.00	401.29	2.2e-14
/cgnl_9/gcgdata/geneseq/geneseq/NA1994.DAT:AX37660				205.00	401.29	2.2e-14
/cgnl_9/gcgdata/geneseq/geneseq/NA1991.DAT:AX37660				201.50	400.08	2.6e-14
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				182.00	363.96	2.7e-12
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				182.00	363.96	2.7e-12
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				164.00	326.19	3.4e-10
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				164.00	322.32	5.5e-10
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				164.00	322.03	5.8e-10
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				164.00	320.95	6.6e-10
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				164.00	320.95	6.6e-10
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AX37660				148.00	286.49	5.5e-08
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AX37660				145.50	287.49	4.8e-08
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AX37660				145.00	286.51	5.5e-08
/cgnl_9/gcgdata/geneseq/geneseq/NA1997.DAT:AX37660				140.50	274.33	2.6e-07
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				121.00	239.85	2.2e-05
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				90.50	163.06	0.4118
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				89.50	174.74	0.0921
/cgnl_9/gcgdata/geneseq/geneseq/NA1998.DAT:AX37660				89.50	166.06	0.2803
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				89.50	166.06	0.2803
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AX37660				89.50	159.47	0.6521
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660				88.00	158.19	0.7690
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				88.00	143.46	9.8e+03
/cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AX37660				87.50	143.64	4.97
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				87.50	141.74	6.34
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660				87.50	141.74	6.34

/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660 + 86.00 153.24 1.45
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660 + 84.50 137.00 11.65
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660 + 83.00 141.49 6.55
/cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AX37660 + 82.50 143.54 5.03
/cgnl_9/gcgdata/geneseq/geneseq/NA1999.DAT:AX37660 + 82.50 115.46 184.49
seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2001.DAT:AAF94480

seq_documentation_block:
ID AAF94480 standard; cDNA; 681 BP.
XX AAF94480;
AC AAF94480;
DT 04-JUN-2001 (first entry)
XX Human hydrophobic domain containing protein clone HP03880 cDNA #104.

XX Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antitumor; osteoparic; anti-inflammatory;
KW cytosolic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response; ss.

XX Homo sapiens.
OS
XX
XX
PN WO2001126660-A2.
XX
PD 22-FEB-2001.
XX
PF 10-AUG-2000; 2000WO-JP05356.
XX

PR 17-AUG-1999; 99JP-0230344.
PR 07-SEP-1999; 99JP-0282551.
PR 01-OCT-1999; 99JP-0281132.
PR 22-OCT-1999; 99JP-0301624.
PR 04-NOV-1999; 99JP-0313877.

XX (SAGA) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
PI WPI; 2001-160059/16.
XX P-PSDB: AAB88590.

XX Human proteins with hydrophobic domains and the DNAs which encode them
XX are useful for treating autoimmune disorders, burns and tumors and for
XX screening novel pharmaceuticals -
XX
XX Claim 3; Page 426-427; 518pp; English.

XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
XX AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
XX anti-HIV, neuroprotective, antianemic, vulnery, antitumor,
XX osteoparic, anti-inflammatory and cytostatic activities, and can be
XX used in gene therapy. (I) can be used as pharmaceuticals and as antigens
XX to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
XX probes for genetic diagnosis and gene sources for gene therapy or for
XX producing (I) in large quantities. Cells containing (II) are used for
XX the detection of ligands or receptors corresponding to membrane or
XX secretory proteins and to screen small molecule novel pharmaceuticals.
XX Antibodies directed to (I) can be used for the detection, quantification
XX and purification of (I). Activities of (I) may include cytokine and cell
XX proliferation/differentiation function, immune stimulating or suppressing
XX activity, haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity and anti-inflammatory
XX activity. (I) and (II) can be used to treat autoimmune disorders e.g.
XX multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
XX inflammatory bowel disease and tumours. (I) and (II) can also be used for
XX wound healing, as nutritional sources or supplements e.g. as amino acid,
XX carbon or nitrogen source, to effect metabolism, catabolism, anabolism,

CC processing and utilisation of dietary fat, protein, carbohydrate, vitamins and minerals, to effect behavioural characteristics, to affect appetite, and can act as antigens in vaccines to raise an immune response to the protein or another material cross-reactive with the protein.

XX

SQ Sequence 681 BP; 179 A; 186 C; 183 G; 133 T; 0 other;

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_23_227 x AAF94480 ..
Align seg 1/1 to: AAF94480 from: 1 to: 681

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
'67 GACGAGATGAGAACACACCCGCTGTGCCCATGAGGCCCTCTTGGACGAGGA 116

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
117 CACCCTCTTTGCCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGAACA 166

34 LeGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
167 TTGGCTGCAAGGTTGTCCTGATTGTAACTACACAGACAGATCACC 216

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
217 TCCTGGATGAGACCGATAGTCAAGTTCCTGGGCGCGTGGACGGCGAAC 266

67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
267 CTATATCTGCTGATGTTGGTATCAGATGCCCTTAGCAGACAGAACCCA 316

84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
317 GACGAGATTCGAGACATTTGGCTGTGTAACAGATATCAAGGGCGCGAC 366

101 LeuLysLysGlyLysIleGlnGlyGlnLeuSerAlaTyrGlnAlaPr 117
|||||
367 CTGAAGAAAGGGAAGATTCAGGGCCAGGAGTTATCAGCCTACCAAGCTCC 416

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
417 CTCGCCACCGGCACACATGCTTCCATCGCTACCAAGTCTTTGCTATC 466

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
467 TTCAGGAAGGAAAGTCATCTCTCTCTCCCAAGGAAACAAACATCGA 516

151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
517 GGCTCTTGAAATGGACAGATTTCTGAACCGTTTCCACCTGGCGCAACC 566

167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
567 TGAACAGACCCAGTTCATGACCCAGACACTACCAGGACTCACCACCC 616

184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
617 TCCAGGCTCCAGAGAAAGGGCCAGGAGCCCAAGCAACAAACACGAGCG 666

201 GluIleAlaAlaCys 205
|||||
667 GAGATAGCTGCCTGC 681

seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX97660

seq_documentation_block:

AAAX97660 standard; DNA; 826 BP.

XX

AC AAX97660;

DT 13-SEP-1999 (first entry)

XX

DE Extended human secreted protein coding sequence, SEQ ID NO. 124.

XX

KW Secreted protein; human; cytokine; cellular proliferation; cell movement;

KW Cellular differentiation; immune system regulator; anti-inflammatory;

KW Hematopoiesis regulator; tissue growth regulator; tumour inhibitor;

KW Reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;

XX genetic disease; ss.

OS Homo sapiens.

XX

PN WO9931236-A2.

XX

PD 24-JUN-1999.

XX

PF 17-DEC-1998; 98WO-IB02122.

XX

PR 10-AUG-1998; 98US-0096116.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0081563.

XX

PA (GEST) GENSET.

XX

PI Bougueleret L, Duclert A, Dumas Milne Edwards J;

DR WPI; 1999-385906/32.

DR P-PSDB; AAY35976.

XX

XX New isolated human secreted proteins

PT

XX Claim 1; Page 255; 516pp; English.

PS

XX This sequence represents an extended human secreted protein coding sequence of the invention. The secreted proteins can be used in treating or controlling a variety of human conditions. The secreted proteins may act as cytokines or may affect cellular proliferation or differentiation or may act as immune system regulators, haematopoiesis regulators, tissue growth regulators, regulators of reproductive hormones or cell movement or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or tumour inhibition activity. The DNAs can be used in forensic procedures to identify individuals or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal expression of the genes corresponding to the extended cDNAs. They are also useful for constructing a high resolution map of the human chromosomes. They can also be used for gene therapy to control or treat genetic diseases.

XX

SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_23_227 x AAX97660 ..
Align seg 1/1 to: AAX97660 from: 1 to: 826

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GACGAGATGAGAACACCCGCTGTGCCCATGAGGCCCTCTTGGACGAGGA 130

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCTCTTTGGCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGGNACA 180

```
34 leCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTGTTCTCTGATTGTAACAACATACAGACAGAAGATCACC 230

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGGACCGGATAGTCAGTTCCTCCGGGGCGGTGGACGGCGCAAC 280

67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCTGGTGTGGTGTGGATCCAGATGCCCTAGCAGACAGAACCCA 330

84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATCTTGGAGACATTTGGTGTAAACAGATATCAAGGGCGCGAC 380

101 LeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
381 CTGAAGAAAGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACGAGCTCC 430

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
431 CTCCTCCAGCCGGCACAGTGGCTTCCTCCGCTACCAAGTTCTTTGTCTATC 480

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArq 150
|||||
481 TTCAGAGAGAAAGGTCACTCTCTCTCTCCCTCCCAAGAGAAACAACTCGA 530

151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
531 GGCTCTTGGAAATGGACAGATTTCTGAACCGTTTCCACCTGGCGGAACC 580

167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAAGACCCAGTTTCATGACCCAGACACTACCAAGGACTCACCAACC 630

184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCCAGGAGCCCAAGCAGCAAAACACAGGCG 680

201 GluIleAlaAlaCys 205
|||||
681 GAGATAGCTGCCTGC 695
```

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT.AAC00013

seq_documentation_block:

```
ID AAC00013 standard; cDNA; 826 BP.
XX
AC AAC00013;
XX
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein cDNA sequence #4.
XX
KW Human; secreted protein; 5' EST; expressed sequence tag; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 15..698
XX /*tag= a
XX /product= "secreted protein"
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
```

```
PA (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
DR P-PSDB; AAG00016.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Example 19; SEQ ID 7; 71pp + CD-ROM; English.
XX
XX The present sequence is a full length cDNA encoding a human
CC secreted protein. The cDNA was obtained from a 5' EST using first and
CC second strand synthesis procedures. 5' ESTs were prepared from total
CC human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region
CC (UTR) of the mRNA because they are often obtained from oligo-dT primed
CC cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC sequences derived from the 5' ends of mRNAs and even in those cases where
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC are also used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. They are used to obtain upstream regulatory sequences
CC and to design expression and secretion vectors.
XX
XX Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;
```

```
alignment_scores:
    Quality: 1115.00      Length: 205
    Ratio: 5.439          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_23_227 x AAC00013
Align seg 1/1 to: AAC00013 from: 1 to: 826
```

```
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GAGGAGGATGAGAACAGCCCGTGTGCCCATGAGGCCCTCTCTGGACGAGGA 130

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCCTCTTTTGGCAGGGCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180

34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTGTTCTCTGATTGTAACAACATACAGACAGAAGATCACC 230

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGGACCGGATAGTCAGTTCCTCCGGGGCGGTGGACGGCGCAAC 280

67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCTCTGGTGTGGATCCAGATGCCCTAGCAGACAGAACCCA 330

84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATCTTGGAGACATTTGGTGTAAACAGATATCAAGGGCGCGCAC 380

101 LeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
381 CTGAAGAAAGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACGAGCTCC 430

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
431 CTCCTCCAGCCGGCACAGTGGCTTCCTCCGCTACCAAGTTCTTTGTCTATC 480
```

CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA24264 and AA24644 to AA24650 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 826 BP; 228 A; 229 C; 211 G; 158 T; 0 other;

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_23_227 x AA242252 ..
Align seg 1/1 to: AA242252 from: 1 to: 826

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GACGAGGATGAGACAGCCCGTGTGCCCATGAGGCCCTCCTGGACGAGGA 130
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCTCTTTGCCAGGGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34 IeGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTTTCTCTGATTGTAACACTACAGACAGACAGATCACC 230
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGAGCCGATAGTCAAGTTCCCGGGCGCGTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCCTGGTGGTGGATCCAGATGCCCTAGCAGAGCAGCAACCCA 330
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATCTGGAGACATTGGCTGGTAACAGATATCAAGGGCGCCGAC 380
101 LeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
381 CTGAGAAAAGGGAAGATTTCAGGGCCAGAGGATTATCAGCCTACCGCTCC 430
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
431 CTCCCCCAGCGCACAGATGGCTTCCATCGCTACCAAGTTCTTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAAGGAAAGGTCATCTCTCTCTTCCCAAGGAAACAAACTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
531 GGCTCTTGGAAAATGGACAGATTCTGAACCGTTTCCACCTGGGGGAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGGACTCACCAACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAAGCACAACCAACAGGCG 680
201 GluIleAlaAlaCys 205
|||||
681 GAGATAGCTGCCTGC 695

seq_name: /cgnl_9/gcgdata/geneseq/NA2000.DAT:AA242252

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAAGGAAAGGTCATCTCTCTTCCCAAGGAAACAAACTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
531 GGCTCTTGGAAAATGGACAGATTCTGAACCGTTTCCACCTGGGGGAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGGACTCACCAACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAAGCACAACCAACAGGCG 680
201 GluIleAlaAlaCys 205
|||||
681 GAGATAGCTGCCTGC 695
seq_name: /cgnl_9/gcgdata/geneseq/NA2000.DAT:AA242252
seq_documentation_block:
ID AA242252 standard; cDNA; 826 BP.
XX
AC
AA242252;
XX
01-FEB-2000 (first entry)
XX
Human phosphatidylethanolamine-binding protein encoding cDNA.
XX
Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification; ss.
XX
Homo sapiens.
XX
W09953051-A2.
XX
21-OCT-1999.
XX
09-APR-1999; 99WO-1B00712.
XX
09-APR-1998; 98US-0057719.
XX
28-APR-1998; 98US-0069047.
XX
(GIST) GENSET.
XX
Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI: 2000-038446/03.
XX
P-PSDB: AA24647.
XX
Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
Example 21; Page 168-169; 837pp; English.
XX
AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AA24651 to
CC AA246538 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the

seq_documentation_block:

ID AAF94490 standard; cDNA; 897 BP.

XX AC AAF94490;

XX DT 04-JUN-2001 (first entry)

XX DE Human hydrophobic domain containing protein clone HP03880 cDNA #114.

XX Human: hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
 KW antianaemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
 KW cytosatic; gene therapy; autoimmune disorder; multiple sclerosis;
 KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
 KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
 KW behavioural characteristic; immune response; ss.

XX OS Homo sapiens.

XX PN WO200112660-A2.

XX PD 22-FEB-2001.

XX PF 10-AUG-2000; 2000WO-JP05356.

XX PR 17-AUG-1999; 99JP-0230344.

XX PR 07-SEP-1999; 99JP-0252551.

XX PR 01-OCT-1999; 99JP-0281132.

XX PR 22-OCT-1999; 99JP-0301624.

XX PR 04-NOV-1999; 99JP-0313877.

XX PA (SAGA) SAGAMI CHEM RES CENT.

XX PA (PROT-) PROTEGENE INC.

XX PI Kato S, Kimura T;

XX WPI: 2001-160059/16.

XX DR P-PSDB; AAB88590.

XX Human proteins with hydrophobic domains and the DNAs which encode them

PT are useful for treating autoimmune disorders, burns and tumors and for
 PT screening novel pharmaceuticals -

XX Claim 4; Page 442-444; 518pp; English.

XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
 CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
 CC anti-HIV, neuroprotective, antianaemic, vulnery, antiulcer,
 CC osteopathic, anti-inflammatory and cytostatic activities, and can be
 CC used in gene therapy. (I) can be used as pharmaceuticals and as antigens
 CC to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
 CC probes for genetic diagnosis and gene sources for gene therapy or for
 CC producing (I) in large quantities. Cells containing (II) are used for
 CC the detection of ligands or receptors corresponding to membrane or
 CC secretory proteins and to screen small molecule novel pharmaceuticals.
 CC Antibodies directed to (I) can be used for the detection, quantification
 CC and purification of (I). Activities of (I) may include cytokine and cell
 CC proliferation/differentiation function, immune stimulating or suppressing
 CC activity, haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
 CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
 CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
 CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
 CC wound healing, as nutritional sources or supplements e.g. as amino acid,
 CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
 CC processing and utilisation of dietary fat, protein, carbohydrate,
 CC vitamins and minerals, to effect behavioural characteristics, to affect
 CC appetite, and can act as antigens in vaccines to raise an immune response
 CC to the protein or another material cross-reactive with the protein.

XX SQ Sequence 897 BP; 226 A; 253 C; 234 G; 184 T; 0 other;

alignment_scores:

Quality: 1115.00 Length: 205
 Ratio: 5.439 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_23_227 x AAF94490 ..

Align seg 1/1 to: AAF94490 from: 1 to: 897

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
 |||||
 165 GAGCAGGATGAGAACGCCGCTGTGCCCATGAGGCCCTCTTGGACGAGGA 214
 |||||
 17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
 |||||
 215 CACCCCTCTTTGCCAGGGCCTTGAAGTTTCTACCCAGAGTTGGGGANCA 264
 |||||
 34 leglyCysLysValValProAspCysAsnTyrArgGlnLysIleThr 50
 |||||
 265 TTGGCTGCAAGGTTGTTCTGATTGTTAAACAATACAGACAGAAAGATCAC 314
 |||||
 51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
 |||||
 315 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGCTGGACGGCGCAAC 364
 |||||
 67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
 |||||
 365 CTATATCCTGGTGGTGGATCCAGATGCCCTTAGCAGAGAGAACCCCA 414
 |||||
 84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
 |||||
 415 GACAGAGATTCTGGAGACATTTGGCTGTAACAGATATCAAGGGCGCGAC 464
 |||||
 101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
 |||||
 465 CTGAAGAAGAGGAAGATTTCAGGGCCAGGAGTTATCAGCCTTACCAGGCTCC 514
 |||||
 117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
 |||||
 515 CTCGCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTGTCTATC 564
 |||||
 134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
 |||||
 565 TTCAGAGAGAAAGATCATCTCTCTCTCCCAAGGAGAAACAAACTCGA 614
 |||||
 151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
 |||||
 615 GGCTCTTGGAAAATGACACAGATTTCGAACCGTTTCCACCTGGGCGAACC 664
 |||||
 167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
 |||||
 665 TGAAGCAAGCACCCAGTTTCATGACCCAGAACTACCAAGACTACCAACCC 714
 |||||
 184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
 |||||
 715 TCACGAGCTCCACAGAAAGGCCAGCGAGGCCCAAGACACAAAACCCAGCG 764
 |||||
 201 GluIleAlaAlaCys 205
 |||||
 765 GAGATAGTGCCTGC 779

seq_name: /cgnl_9/gcdata/geneseq/geneseq/NA2001.DAT: AAC84882

seq_documentation_block:

ID AAC84882 standard; cDNA; 932 BP.

XX AC AAC84882;

XX DT 20-APR-2001 (first entry)

XX DE Human SECl nucleic acid sequence (clone ID 3445452).

Thu Nov 15 10:52:14 2001

us-09-215-435-225_copy_23_227.rng

XX SECX; cytostatic; gynecological; gene therapy; screening assay; human;
KW chromosomal mapping; forensic biology; cell proliferation; cancer;
KW cell differentiation; immune associated disorder; gestational disease;
KW SEC1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 113..781
FT /*tag= a
FT /product= "SEC1"
FT 113..181
FT sig_peptide /*tag= b
FT mat_peptide 182..778
FT /*tag= c
XX
XX WO200078802-A2.
XX
XX
XX 28-DEC-2000.
XX
XX 23-JUN-2000; 2000WO-US17328.
XX
XX 23-JUN-1999; 99US-0140584.
XX 20-JUL-1999; 99US-0144722.
XX 16-SEP-1999; 99US-0154520.
XX 22-JUN-2000; 2000US-0604286.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
XX Herrmann JL;
XX
XX WPI: 2001-071385/08.
XX P-PSDB; AAB48368.
XX
XX Polynucleotides encoding SECX proteins useful for treating disease
XX characterized by an aberrant level of cell proliferation and/or
XX differentiation like cancer or immune associated disorders -
XX
XX Claim 3; Fig 1; 132pp; English.
XX
XX The invention relates to human SECX polypeptides and polynucleotides
XX encoding them. The SECX polypeptides can be expressed by standard
XX recombinant methodology. The SECX polypeptides are useful for treating
XX or preventing a SECX-associated disorder. The invention is useful in
XX screening assays; detection assays (e.g. chromosomal mapping, cell and
XX tissue typing, forensic biology); predictive medicine (diagnostic assays,
XX prognostic assays, monitoring clinical trials, and pharmacogenomics); and
XX methods of treatment (e.g. therapeutic and prophylactic), especially
XX disorders characterized by aberrant cell proliferation and/or
XX differentiation like cancer or immune associated disorders or gestational
XX disease. The present sequence represents a SEC1 nucleic acid sequence.
XX
XX Sequence 932 BP; 244 A; 257 C; 244 G; 187 T; 0 other;

alignment_scores:
Quality: 1111.00 Length: 205
Ratio: 5.420 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.512

alignment_block:

US-09-215-435-225_COPY_23_227 x AAC84882 ..

Align seg 1/1 to: AAC84882 from: 1 to: 932

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17

179 GACGAGGATGAGAACACCGCGTGTGCCCATGAGGCCCTCTTGGACGAGGA 228

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34

|||||

229 CACCCTCTTTTGGCAGGGCCTTGAAGCTTTTCTACCCAGAGTTGGGAACA 278
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
279 TTGGCTCAAGGTGTCTCTGATTGTAACAACATACAGACAGAAGATCACC 328
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
329 TCCGTGGATGGAGCGGATAGTCAAGTTCCCGGGGCGCTGGACGGCGAAC 378
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
379 CTATATCTGTGTATGTTGGATCCAGATCCCTAGCAGACAGAACCCCA 428
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
429 GACAGAGATTCTGGAGACATTGGCTGGTACAGATATCAAGGGCGCGAC 478
101 LeuLysLysGlyLysIleGlnGlyGlnLeuSerAlaTyrGlnAlaPr 117
|||||
479 CTGAAGGAAGGGAAGATTACGGGCCAGGAGTTATCAGCCTACCAAGGCTCC 528
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
529 CTCCTCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTGTCTATC 578
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
579 TTCAGGAAGGAAAGTCATCTCTCTCTCCCAAGGAAACAAACTCGA 628
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
629 GGCTCTCTGGAATAATGGACAGATTTCTGAACCGTTTCCACCTGGGGGAACC 678
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
679 TGAGACAGACACCCAGTTCATGACCCAGCACTACCAAGGACTCACCAACCC 728
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
729 TCCAGGCTCCCAAGAGAGGCGCCAGCAGCCCAAGCACAACCAACCGCG 778
201 GluIleAlaAlaCys 205
|||||
779 GAGATAGCTGCCTGC 793

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/nA2000.DAT.AAA78426

seq_documentation_block:

ID AAA78426 standard; cDNA; 1028 BP.

XX AAA78426;

XX AC

XX 20-NOV-2000 (first entry)

XX DT

XX DE Human secreted protein gene 46 SEQ ID NO:56.

XX KW Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW antinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
KW neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW immune system disorder; angiogenesis; hyperproliferative disorder;
KW cardiovascular disorder; apoptosis; neurological disease;
KW infectious disease; wound healing; ss.

XX OS Homo sapiens.

XX XX

XX WO200035937-A1.

XX XX

XX PD 22-JUN-2000.

XX XX

XX PF 16-DEC-1999; 99WO-US29950.

XX XX

PR 17-DEC-1998; 98US-0112809.
PR 18-DEC-1998; 98US-0113006.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX WPI; 2000-431566/37.
DR P-PSDB; AAB24482.
XX
XX Forty seven human nucleic acids encoding secreted proteins, useful in
PT the treatment, prevention and diagnosis of cancers, disorders of the
PT immune system, angiogenesis disorders, neurological diseases and
PT hyperproliferative disorders -
XX
XX Claim 1; Page 475; 562pp; English.
XX
XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the
CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; antianaemic;
CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
CC antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
CC nootropic; neuroprotective; antimicrobial and antiparkinsonian.
CC Human secreted protein polynucleotides, polypeptides, antagonists and/or
CC agonists may be useful in treating, preventing, and/or diagnosing other
CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
CC disorders of the immune system; (c) angiogenesis disorders; (d)
CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
CC associated with increase apoptosis; (g) neurological diseases; and
CC (h) infectious diseases. They are also used to promote wound healing.
CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 1028 BP; 309 A; 268 C; 259 G; 186 T; 6 other;

alignment_scores:
Quality: 1109.00 Length: 205
Ratio: 5.436 Gaps: 0
Percent Similarity: 99.512 Percent Identity: 99.512

alignment_block:
US-09-215-435-225_COPY_23_227 x AAA78426 ..
Align seg 1/1 to: AAA78426 from: 1 to: 1028

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
219 GAGCAGGATGAGAACAGCCGCTGTGCCATGAGGCCCTCTTGGACGAGGA 268

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
269 CACCCCTCTTTGCCAGGGCTTGAAGTTTCTACCCAGAGTTGGGAACA 318

34 leGlyCysLysValValProAspCysAsnTyrArgGlnLysIleThr 50
|||||
319 TTGGCTGCAGGTTGTTCTGATTGTAACTACACACAGACAGATCACC 368

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
369 TCCTGGATGAGCGGATAGTCAAGTTCCCGGGGCGGTGGACGGCGCAAC 418

67 rTyrlleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
419 CTATATCCTGGTGGATGGATCCAGATGCCCTAGCAGACAGAACCCA 468

84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
469 GACAGAGATCTGGAGACATATGCTGGTAAACAGATATCAAGGGCGCCGAC 518

101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117

|||||
519 CTGAGAGRAAGGAAGATTTCAGGGCCAGAGATTATCAGCTTACCAGGCTCC 568
|||||
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
569 CTCCTCCACCGGCACACAGTGGCTTCATCGCTACACAGTCTTTGTCTATC 618
|||||
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
619 TTCAGGAAGAGAAAGTCATCTCTCTCTCCCAAGCAAAACAACTCGA 668
|||||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
669 GGCCTCTGTGAAAATGACAGATTCTGAACCGCTTCCACCTGGGCGAACC 718
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
719 TGAAGCAACACCCAGTTCATGACCCAGAACTACACAGGACTCACCACCC 768
|||||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
769 TCCAGGCTCCACAGAAAGGGCCAGCGAGCCCAAGCACACAAACACGAGCG 818
|||||
201 GluIleAlaAlaCys 205
|||||
819 GAGATAGCTGCCTGCG 833

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT: AAC77539
seq_documentation_block:
ID AAC77539 standard; cDNA; 909 BP.
XX AAC77539;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORF3094 polynucleotide sequence SEQ ID NO:6187.
XX
XX Human; open reading frame; OREF; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
OS
XX WO200058473-A2.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF
XX
XX 31-MAR-1999; 99US-0127607.
PR
XX 02-APR-1999; 99US-0127636.
PR
XX 05-APR-1999; 99US-0127728.
PR
XX 30-MAR-2000; 2000US-0540763.
PR
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shimkets RA, Leach M;
PI
XX WPI; 2000-602362/57.
XX
XX P-PSDB; AAB43330.
XX

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
566 TTCAGAGAGAAAGTCATCTCTCTCTCCCAAGGAAAAACAACTCGA 615
|||||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyClnPr 167
|||||
616 GGCTCTTGGAAAAATGCACAGATTCTGAAACCGTTTCCACCTGGCGAACC 665
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
666 TGAAGCAAGCACCCAGTTCATGACCCAGAACTACCAAGGACTCACCACCC 715
|||||
184 euGlnAlaProArgGluAArgAlaSerGlnProLysHisLysAsnGlnAla 200
|||||
716 TCCAGGCTCCAGAGAAAGGCCAGCGAGCCCAAGCACAAAAACCAAGCG 765
|||||
201 GluIleAlaAlaCys 205
|||||
766 GAGATAGCTGCCTGC 780
|||||
seq_name: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA15582
seq_documentation_block:
ID AAA15582 standard; cDNA; 903 BP.
XX
XX
XX
XX
01-AUG-2000 (first entry)
DT
XX

Human: phospholipid binding protein; PLBP2; foetal development disorder;
reproduction disorder; cell proliferation disorder; immune response;
autoimmune disorder; AIDS; infertility; cytostatic; immunosuppressive;
gene therapy; hereditary neuropathy;
phosphatidylethanolamine binding protein D1; PE-BP D1; ss.
Homo sapiens.

alignment_block:
rs-00-215-225 COPY 23 227 x AAC775539

Key
cnc

```

Key      Location/Qualifiers
CDS      88..771
          /*tag= a
          /product= "Human PLBP2"

US0603767-A.

16-MAY-2000.

09-DEC-1998; 98US-0208718.

28-OCT-1997; 97US-0958820.

{INCY-) INCYTE PHARM INC.

Corley NC, Shah P, Lal P, Hillman JU
WPI; 2000-375529/32.
P-PSDB; AAY94263.

New purified phospholipid binding prote
diagnosing, treating or preventing dise
fetal development, reproduction, cell F
response -

Example 5: Fig 2; 37pp; English.

```

Example 5: Fig 2; 37pp; English.

XX The present sequence is the phospholipid binding protein 2 (PLBP2) gene.
CC This gene is expressed in lung, prostate and heart tissues. Also, the
CC protein is expressed in foetal tumour tissues. PLBP2 may be used for the
CC diagnosis, prevention, or treatment of disorders associated with foetal
CC development (e.g. hereditary neuropathies), reproduction (e.g.
CC infertility), cell proliferation (e.g. cancers), and the immune response
CC (e.g. PLBP2 antibodies may also be developed for potential drug

516 CTCCTCCACCCGGGCACACACAGTGGCTTCCATCGCTACCACTTCTTGTCTATC 503

Thu Nov 15 10:52:14 2001

```
1  AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
129 GACGAGGATGAGAACAGCCGCTGCCCCATGAGGCCCTCTTGACGAGGA 178
17  pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
179 CACCCCTCTTTGCCAGGCGCTGGAAGTTTCTACCCAGAGTTGGGGAACA 228
34  leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
229 TTGGCTGCAAGGTGTTCTCTGATGTTGAACACTACACAGAGAATCACC 278
51  SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaT 67
279 TCTCGGATGGAAGCCAGTAGTCAAGTTCCCGGGGCGCTGGACGGCGCAA 328
67  hrTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluPr 83
329 CCTATAATCTGGTGATGTGGATCCAGATGCCCTAGCAGAGCAGAACC 378
83  oArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaA 100
379 CAGACAGAGATTCGGAGACATGGCTGGTGAACAGATATCAAGGGCGCCG 428
100 spLeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAla 116
429 ACCTGAGAAAGGAAGATTCAGGGCCAGGAGTTATCAGCCTACCAAGCT 478
117 ProSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTy 133
479 CCCTCCCGCCAGCGCACACAGTGGCTTCATCGCTACCAAGTCTTTGCTA 528
133 rLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrA 150
529 TCTTCAGGAAGAAAGATCATCTCTCTCCATCCAGGAAACAAACTC 578
150 rgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGlu 166
579 GAGGCTCTTGGAAATGACAGATTTCTGAACCGTTTCCACCTGGGGGAA 628
167 ProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProTh 183
629 CCTGAAGCAAGCACCCAGTTTCATGCCAGAACTACCAGGACTCACCAAC 678
183 rLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnA 200
679 CTCCAGGCTCCAGAGAAAGGGCCAGCGAGCCCAAGCAGCAACAAACAGG 728
200 laGluIleAlaAlaCys 205
729 CGGAGATAGCTGCCTGC 745
seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA2000.DAT:AAA96350
seq_documentation_block:
ID  AAA96350 standard; cdNA; 890 BP.
XX
AC  AAA96350;
XX
DT  08-FEB-2001 (first entry)
XX
DE  cDNA encoding a novel polypeptide designated PRO4408.
XX
KW  Secreted protein; transmembrane protein; PRO1484; PRO1122;
KW  PRO1889; PRO1890; PRO1785; PRO4357; PRO4405; PRO4356;
KW  PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990;
KW  PRO4424; PRO4422; PRO4430; PRO4499; PRO1889 polypeptide is
KW  insulinemia; kidney disorder; Bergers disease; nephropathy;
KW  Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis;
KW  Crohns disease; ss.
XX
OS  Homo sapiens.
```

```
XX FH Key Location/Qualifiers
FT CDS 89..760
FT FT /*tag= a
FT FT 89..154
FT FT sig_peptide /*tag= b
XX
XX WO200056889-A2.
XX
XX 28-SEP-2000.
XX
XX 01-MAR-2000; 2000WO-US05601.
XX
XX 23-MAR-1999; 99US-0125774.
XX 23-MAR-1999; 99US-0125778.
XX 24-MAR-1999; 99US-0125826.
XX 31-MAR-1999; 99US-0127035.
XX 05-APR-1999; 99US-0127706.
XX 21-APR-1999; 99US-0130359.
XX 27-APR-1999; 99US-0131270.
XX 27-APR-1999; 99US-0131272.
XX 27-APR-1999; 99US-0131291.
XX 04-MAY-1999; 99US-0132371.
XX 04-MAY-1999; 99US-0132379.
XX 04-MAY-1999; 99US-0132383.
XX 25-MAY-1999; 99US-0135750.
XX 08-JUN-1999; 99US-0138166.
XX 20-JUL-1999; 99US-0144791.
XX 03-AUG-1999; 99US-0146070.
XX 09-DEC-1999; 99US-0170262.
XX
XX (GETH ) GENENTECH INC.
XX
XX Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
XX Stewart RA, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2000-628263/60.
XX
XX P-PSDB; AAB18923.
XX
XX Novel secreted and transmembrane polypeptides useful for diagnosing
XX tumour in a mammal, for identifying agonists and antagonists of the
XX polypeptide and for therapeutic use
XX
XX Claim 2; Fig 29; 222pp; English.
XX
XX The present sequence encodes a secreted or transmembrane polypeptide.
XX The specification describes polypeptides designated PRO1484, PRO4334,
XX PRO1122, PRO1889, PRO1890, PRO1785, PRO4353, PRO4357, PRO4405,
XX PRO4356, PRO4380, PRO4352, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,
XX PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is
XX useful for diagnosing tumour in a mammal. The polypeptides, their
XX agonists and antagonists are useful treating a condition associated with
XX expression or activity of the polypeptide. Conditions treated include
XX obesity, diabetes or hyper-or hypo-insulinemia. The polypeptides are
XX capable of inducing proliferation of mammalian kidney mesangial cells
XX and are therefore useful for treating kidney disorders associated with
XX decreased mesangial cell function such as Bergers disease or other
XX nephropathies associated with Schonlein-Henoch purpura, celiac disease,
XX dermatitis herpetiformis or Crohns disease. The nucleic acids may be used
XX to generate transgenic animals for use in development and screening of
XX therapeutically useful reagents and also for chromosome identification
XX and tissue typing.
XX
XX Sequence 890 BP; 228 A; 246 C; 234 G; 182 T; 0 other;
XX
XX alignment_scores:
XX Quality: 1080.00 Length: 205
XX Ratio: 5.294 Gaps: 0
XX Percent Similarity: 99.512 Percent Identity: 99.512
XX
XX alignment_block:
XX US-09-215-435-225_COPY_23_227 x AAA96350
```


84 rgGlnhrgPheTrpArgHisTrrpleuValThrAspIleLysGlyAla 99
|||||
401 GACAGAGATTCTGGAGACATTCGGCTGGTAACAGATATCAAGGGCC 447

seq_name: /cgnl_9/gcgdata/geneseq/geneseq/NA1992.DAT:AAQ30001
seq_documentation_block:
ID AAQ30001 standard; cDNA to mRNA; 1047 BP.
XX
XX
XX AC AC
XX
XX DT DT
XX
XX DE DE
XX HCNP precursor gene #1.
XX
XX Hippocampal cholinergic neurotrophic peptide; HCNP; hippocampal tissue;
KW neonatal rat; expression vector; neurodegenerative disorder; dementia;
KW Alzheimer's disease; Parkinson's disease; ss.
XX
XX Rattus norvegicus.
OS

Key	Location/Qualifiers
CDS	26..589
5'UTR	/*tag= a
3'UTR	1..25
polyA_signal	587..1047
misc_RNA	/*tag= b
	987..992
	/*tag= c
	29..61
	/*tag= d
	/*tag= e
	/label= Neurotrophic_activity_region

EP511816-A.
04-NOV-1992.
27-APR-1992; 92EP-0303800.
27-APR-1991; 91JP-0124688.
(OJIK/) OJIK K.
(SUMU) SUMITOMO PHARM CO LTD.
(YAMA/) YAMAMOTO M.
Aqui H, Fukushima N, Irie T, Kojima S, Nishihara T;
Ojika K, Ono K, Tohdoh N, Tojo S, Ueki Y;
WPI; 1992-367633/45.
P-PSDB; AAR27897.
Neurotrophic peptide derivs. - used for treating neurological
degenerative disorders, e.g. Alzheimer's disease or Parkinson's
disease
Claim 3; Page 31-32; 57pp; English.
The sequence given is the hippocampal cholinergic neurotrophic peptide
(HCNP) gene. The active part of the protein encoded by this sequence
is located in the first eleven amino acids at the N terminus. This
gene was obtained by preparing mRNA from the hippocampal tissue of
neonatal rats, 12 days after birth. The mRNA is converted to cDNA by
standard methods. The cDNA can be used in expression vectors to
produce both the full length protein and the active peptide fragment.
The gene and the peptide are useful for the treatment of
neurodegenerative disorders and dementia eg. Alzheimer's and
Parkinson's disease.
Sequence 1047 BP; 257 A; 296 C; 290 G; 204 T; 0 other;

xx This sequence encodes the human phosphatidylethanolamine binding
 CC protein, hPEBP. This cDNA was isolated from a cDNA library prepared
 CC using RNA extracted from the human glioblastoma cell line T98G. The
 CC protein can be used for the prevention of or in the treatment of
 CC aplasia or abnormal proliferation of glia, neurons or hematopoietic
 CC cells, depression or enhancement of immunological or neurological
 CC activity, inflammatory disease, tumours or diseases induced by
 CC abnormal lipid metabolism. Anti-PEBP antibodies may be used for
 CC diagnosis of genetic diseases.

xx
 SQ Sequence 1001 BP: 250 A; 249 C; 279 G; 223 T; 0 other;

alignment_scores: Quality: 205.00 Length: 151
 Ratio: 2.228 Gaps: 4
 Percent Similarity: 60.927 Percent Identity: 33.113

alignment_block:
 US-09-215-435-225_COPY_23_227 x AAQ80734 ..
 Align seg 1/1 to: AAQ80734 from: 1 to: 1001

30 GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 46
 ||||| : ||
 158 GAGCTGGGCAAGTG.....CTGACCCCCACCCAGGTTAAGAATAG 198
 ||||| : ||
 46 gGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAlav 63
 ||||| : ||
 199 ACCCACCAGCATTCGTGG.....GATGGTCTTG 227
 ||||| : ||
 63 alAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 79
 : : : : : : :
 228 ATTCAGGGAAGCTCTACACCTTGGTCTGTGACAGACCCGGATGCTCCAGC 277
 ||||| : ||
 80 ArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspI1 96
 ||||| : ||
 278 AGGAAGGATCCCAATACAGAAATGGCATCATTTCTGTGGTCAACAT 327
 : : : : : : :
 96 eLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSera 113
 : : : : : : :
 328 GAAGGCAATGACATCAGCAGTGGCAGATC.....CTCTCCG 365
 ||||| : ||
 113 laTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGln 129
 ||||| : ||
 366 ATTATGTGGGCTCGGGGCTCCCAAGGGCAGAGCCCTCCACCGCTATGTC 415
 : : : : : : :
 130 PhePheValTyrLeuGlnGluGlyLysValIleSer.....Le 142
 : : : : : : :
 416 TGGCTGGTTTACGACGACGAGCCGCTAAAGTGTGACGAGCCCATCCT 465
 ||||| : ||
 142 uLeuProLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheL 159
 : : : : : : :
 466 CAGCAACCGATCTGGAGACCAACCGTGGCAATTCAGGTGGCGCTCCTCC 515
 ||||| : ||
 159 euAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThr 175
 : : : : : : :
 516 GTAAAAAGTATGAGCTCAGGGCCCGGCTGGCTGGCACGCTTACCAGGCC 565
 ||||| : ||
 176 Gln 176
 : :
 566 GAG 568

OM of: US-09-215-435-225_COPY_23_227 to: Issued_Patents_NA:* out_format : pfs

Date: Nov 15, 2001 5:03 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O/cgnl_1/USPTO.spool/US09215435/runat_14112001_141104_2141/app_query.fasta_1.1519
-DB-issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09215435_@cgnl_1_161 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_23_227

Query length: 205

Database: Issued_Patents_NA.*

Database sequences: 351203

Database length: 113238999

Search time (sec): 228.030000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4	+	1104.00	2458.21	2.8e-129	903
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-403-378B-3	+	212.50	448.08	2.6e-17	1047
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-403-378B-14	+	205.00	427.44	3.7e-16	1447
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-07-644-372-1	+	201.50	426.10	4.4e-16	822
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-1	+	164.00	341.17	2.4e-11	855
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-3	+	164.00	341.15	2.4e-11	856
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-998-416-1021	-	112.50	249.57	3.0e-06	726
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-998-416-991	+	115.00	233.23	2.4e-05	692
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-2	+	105.00	109.29	194.62	4403765
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-103-840A-1	+	105.00	109.26	195.13	4411529
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-47	+	87.50	171.14	0.0698	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-11	+	87.50	171.14	0.0698	700
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-082-614A-47	+	87.50	171.14	0.0698	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-8	+	87.50	148.86	1.22	4794
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-45	+	87.50	148.86	1.22	4795
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-09-082-614A-45	+	87.50	159.88	0.2960	700
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-082-614A-46	+	82.50	159.88	0.2960	700
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-459-499-10	+	82.00	158.77	0.3413	699
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-07-765-830A-5	+	80.00	133.34	8.90	4258
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-056-105-10	+	77.50	126.09	22.54	4895
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-056-105-9	+	77.00	116.21	80.10	11495
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-288-630-1	+	77.00	140.03	3.77	1333
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-08-814-095-7	-	77.00	132.17	484.84	35060
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-697-610-1	+	76.00	131.76	10.89	2240
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-349-357-1	+	76.00	131.76	10.89	2240
/cgnl_7/ptodata/1/ina/PCPUS_COMB.seq:US-09-16980-1	+	76.00	131.16	11.76	2359
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-211-417-2	+	76.00	127.33	19.24	3286
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-770-379-19	+	76.00	99.90	648.37	35100
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-757-669A-19	+	75.00	99.90	648.37	35100
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-791-1158-8	-	75.00	136.20	6.17	1257
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-368-590-1	-	74.00	125.53	24.24	3160
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-812-13	+	74.00	110.51	124.09	7812
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-812-13	+	74.00	110.51	166.31	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-920-812-13	+	74.00	110.51	166.31	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-921-177-13	+	74.00	110.51	166.31	9515
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-362-572-13	+	74.00	110.51	166.31	9515
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-920-828-13	+	74.00	110.51	166.31	9515
/cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-492-027A-9	+	73.50	127.83	18.05	1935
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-785-241-2	+	73.50	122.63	35.14	3031

/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-189-760-5 + 73.00 129.43 14.70 1529
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-188-811-5 + 73.00 129.43 14.70 1529
/cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-514-422-5 + 73.00 129.43 14.70 1529
/cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-656-984A-33 + 73.00 128.29 17.01 1687
/cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-188-811-1 + 73.00 128.17 17.26 1704
seq_name: /cgnl_7/ptodata/1/ina/6A_COMB.seq:US-09-208-718-4

seq_documentation_block:
; Sequence 4, Application US/09208718
; Patent No. 6063767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PHOSPHOLIPID BINDING PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0379 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGPUT12
; CLONE: 3126479
; US-09-208-718-4

alignment_scores:
Quality: 1104.00 Length: 205
Ratio: 5.412 Gaps: 0
Percent Similarity: 99.512 Percent Identity: 99.024
alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-208-718-4 ..
Align seg 1/1 to: US-09-208-718-4 from: 1 to: 903
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
154 GAGCAGATGAGAACAGCCCGTGCCATGAGCCCTCTTGGACGAGGA 203
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
204 CACCCCTCTTTTGGCAGGCGCTTGAAGTTTTTCTACCCAGAGTTGGGGAACA 253

Thu Nov 15 10:52:15 2001

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/403,378B
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 3-124688
 FILING DATE: 27-APR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 1-080398
 FILING DATE: 30-MAR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 1-280590
 FILING DATE: 27-OCT-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 1-333241
 FILING DATE: 21-DEC-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-243003
 FILING DATE: 12-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/758,043
 FILING DATE: 12-SEP-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/873,764
 FILING DATE: 27-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP93/01214
 FILING DATE: 27-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGART, WADDELL A
 REGISTRATION NUMBER: 24,861
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-7860
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1047 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Rattus norvegicus
 STRAIN: Wistar
 TISSUE TYPE: hippocampal tissue of brain
 IMMEDIATE SOURCE:
 CLONE: A010-12
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 26...586
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..25
 FEATURE:
 NAME/KEY: mat.peptide
 LOCATION: 26...586
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 587..1047
 FEATURE:
 NAME/KEY: polyA_site
 LOCATION: 1008..1047
 FEATURE:
 NAME/KEY: polyA_signal
 LOCATION: 987..992

US-08-403-378B-3

34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
 254 TTGGCTGCAAGGTGTCTCTGATGTAACTACAGACAGAGATCACC 303
 51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
 304 TCCTGGATGAGCGGATAGTCAAGTTCCCGGGGCGCTGGACGGCGCAAC 353
 67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
 354 CTATATCTGTGTGTGTGATCCAGATGCCCTAGCAGAGCAAGACCCA 403
 84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
 404 GACAGAGATTCTGGAGACATTGGCTGTACACATATCAAGGGCGCGGAC 453
 101 LeuLysLysGlyLysIleGlnGlnGlnLeuLeuSerAlaTyrGlnAlaPr 117
 454 CTGAAGGAAGGAAGATTCAGGGCCAGAGATTATCAGCCTTACAGGCTCC 503
 117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
 504 CTCGCCACCGCACACAGTGGCTTCATCGCTACCAAGTTCTTTGTCTATC 553
 134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
 554 TTCAGGAAGGAAGAAATCTCTCTCTCCCAAGGAAACAAACATCGA 603
 151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
 604 GGCTCTTGGAAATGGACAGATTCTGAACCGCTTCCACCTGGCGGACCC 653
 167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
 654 TCAAGCAAGCACCAGTTCATGATCCAGCAACTACCAAGGACTCACCACCC 703
 184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
 704 TCCAGGCTCCAGAGGAAGGGCCAGCGAGCCCAAGCAACAAACACCGCG 753
 201 GluIleAlaAlaCys 205
 754 GAGATAGTGCCTGC 768

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:US-08-403-378B-3

seq_documentation_block:
 ; Sequence 3, Application US/08403378B
 ; Patent No. 5759991
 ; GENERAL INFORMATION:
 ; APPLICANT: TOHDOH, NAKOI
 ; APPLICANT: TOJO, SHIN-ICHIRO
 ; APPLICANT: KOJIMA, SHIN-ICHI
 ; APPLICANT: UEKI, YASUYUKI
 ; APPLICANT: NISHIHARA, TOSHIO
 ; APPLICANT: FUKUSHIMA, NOBUYUKI
 ; APPLICANT: IRIE, TSUNEMASA
 ; APPLICANT: ONO, KEIICHI
 ; APPLICANT: AGUI, HIDEO
 ; APPLICANT: OJIKI, KOSEI
 ; TITLE OF INVENTION: NEUROTROPHIC PEPTIDE DERIVATIVES
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
 ; STREET: 2100 PENNSYLVANIA AVENUE, NW
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20037-3202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

```
alignment_scores:
  Quality: 212.50      Length: 186
  Ratio: 2.063        Gaps: 8
  Percent Similarity: 55.376      Percent Identity: 32.796

alignment_block:
US-09-215-435-225_COPY_23_227 x US-08-403-378B-3 ..
Align seg 1/1 to: US-08-403-378B-3 from: 1 to: 1047

30  GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
131 GAGCTGGGCAAGTG.....CTGACGCCACCCAGGTCATGTAATAG 171

46  gGlnLysIleThrSerTyrMetCluProIleValLysPheProGlyAlaV 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 ACCAGCAGCATTTCATGG.....GATGGCCTTG 200

63  aIAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 ATCTGGGAAGCTCTACACCTGGTCTCACAGACCCCGATGCTCCACG 250

80  ArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspI 96
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 AGGAGGACCCCAAAATTCAGGAGTGGCACCACCTCTCTGGTGTCAACAT 300

96  eLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnLeuSerA 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 GAAGGCCAACACATTAGCAGTGGCACTGTC.....CTCTCG 338

113 latYrGlnAlaProSerProAlaHisSerGlyPheHis...ArgTyr 128
|||:|||||:|||||:|||||:|||||:|||||:|||||
339 AATAC...GTGGCTCGGACCTAAAGACACAGGCTCTGCCCGCCCTAC 385

129 GlnPheValTyrLeuGlnGluGlyLysValIleSer..... 141
|||||:|||||:|||||:|||||:|||||:|||||
386 GTCTGCTGCTGTATGACGAGCAGGACGCTCTCAACTGTGACGAGCCCAT 435

142 .LeuLeuProLysGluAsnLysThrArgLysSerTrpLysMetAspArg 158
|||:|||||:|||||:|||||:|||||:|||||:|||||
436 CCTCAGCAACAAGTCCTGGAGACCGCGGCAAGTTCAAGGTGAGTCCT 485

158 heLeuAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMet 174
||:|||||:|||||:|||||:|||||:|||||:|||||
486 TCCGCAAGAAAGTACCCTGGGAGCCCGGTGGCCGCGACGCTTCCAG 535

175 ThrGlnAsnTyrGlnAspSer..... 181
|||||:|||||:|||||:|||||:|||||:|||||
536 GCAGAG...TGGGATGACTCTGTGCCCAAGCTGCATGATCAGTCGGCTG 582

182 .....ProThrLeuGlnA 186
|||||:|||||:|||||:|||||:|||||
583 GAAGTAGGGGCGGTGACGAGCCGCGAGCCCGGGGACCCACAGTACAGT 632

186 laProArg 188
||:|||||:|||||:|||||:|||||:|||||
633 CRAAGTCGT 640

seq_name: /cgnl_7/ptodata/1/ina/5A_COMB.seq:us-08-403-378B-14

seq_documentation_block:
; Sequence 14, Application US/08403378B
; Patent No. 5759991
; GENERAL INFORMATION:
; APPLICANT: TOHDOH, NAOKI
; APPLICANT: TOJO, SHIN-ICHIRO
; APPLICANT: KOJIMA, SHIN-ICHI
; APPLICANT: UEKI, YASUYUKI
; APPLICANT: NISHIHARA, TOSHIO
; APPLICANT: FUKUSHIMA, NOBUYUKI
; APPLICANT: IRIE, TSUNEMASA
; APPLICANT: ONO, KEIICHI
```

```
; APPLICANT: AGUI, HIDEO
; APPLICANT: OJIKAWA, KOSEI
; TITLE OF INVENTION: NEUTROPHILIC PEPTIDE DERIVATIVES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MIION, ZINN, MACPEAK & SEAS
; STREET: 2100 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,378B
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-124688
; FILING DATE: 27-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1-080398
; FILING DATE: 30-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1-280590
; FILING DATE: 27-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1-333241
; FILING DATE: 21-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-243003
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/758,043
; FILING DATE: 12-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/873,764
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP93/01214
; FILING DATE: 27-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGART, WADDELL A
; REGISTRATION NUMBER: 24,861
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: placental tissue
; IMMEDIATE SOURCE:
; CLONE: pl-3
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 120..680
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..119
; FEATURE:
```

Thu Nov 15 10:52:15 2001

us-09-215-435-225_copy_23_227.rni

```
; NAME/KEY: mat_peptide
; LOCATION: 120...680
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 681..1447
US-08-403-378B-14

alignment_scores:
  Quality: 205.00      Length: 151
  Ratio: 2.228         Gaps: 4
  Percent Similarity: 60.927  Percent Identity: 33.113

alignment_block:
US-09-215-435-225_COPY_23_227 x US-08-403-378B-14 ..
Align seg 1/1 to: US-08-403-378B-14 from: 1 to: 1447

30 GluLeuGlyAsnIleGlyCysLysValProAspCysAsnAsnTyrAr 46
|||||
225 GAGCTGGGCAAGTG.....CTGACGCCACCACCGAGGTAAAGATAG 265

46 gClnIysIleThrSerTrpMetGluProIleValLysPheProGlyAlav 63
|||||
266 ACCCACCAGCATTCGTGG.....GATGGCTTG 294

63 alAspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSer 79
|||||
295 ATTCAGGGAAGCTCTACACCTTGCTGCTGACAGACCGCGATGCTCCAGC 344

80 ArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAsp11 96
|||||
345 AGGAAGGATCCCAATACAGATGCGATCATTTCTGTGGTGGTCAACAT 394

96 eIysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnLeuSerA 113
|||||
395 GAAGGCAATGACATCAGCATGGGCACAGTC.....CTCTCCG 432

113 laTyrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGln 129
|||||
433 ATTATGTGGCTCGGGGCTCCCAAGGACACAGGCTCCACCGCTATGTC 482

130 PhePheValTyrLeuGlnGluLysValIleSer.....Le 142
|||||
483 TGGCTGGTTTACGACGACGACAGCGCGTAAAGTGTGACGAGCCCATCT 532

142 uLeuProLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheL 159
|||||
533 CAGCAACCGATCTGGAGACCACTGGTGGCAATTCAGGTGGCTCTTCC 582

159 euAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThr 175
|||||
583 GTAAAAAGTATGAGCTAGGCGCCCGGTGGTGGCCACGTGTATCCAGGCC 632

176 Gln 176
633 GAG 635

seq_name: /cgn1_7/ptodata/1/lna/5A_COMB.seq:US-07-644-372-1

seq_documentation_block:
; Sequence 1, Application US/07644372
; Patent No. 5416009
; GENERAL INFORMATION:
; APPLICANT: Lazzeri, Mario E.
; APPLICANT: Nutman, Thomas B.
; APPLICANT: Weiss, Nikolaus
; TITLE OF INVENTION: A DNA SEGMENT ENCODING A SPECIFIC
; TITLE OF INVENTION: IMMUNODIAGNOSTIC ANTIGEN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
```

```
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/644,372
; FILING DATE: 19910123
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 52..507
US-07-644-372-1
```

```
alignment_scores:
  Quality: 201.50      Length: 147
  Ratio: 2.144         Gaps: 7
  Percent Similarity: 63.946  Percent Identity: 36.054
```

```
alignment_block:
US-09-215-435-225_COPY_23_227 x US-07-644-372-1 ..
```

```
Align seg 1/1 to: US-07-644-372-1 from: 1 to: 822
```

```
33 AsnIleGlyCysLysValValPro...AspCysAsnAsnTyrArgGlnly 48
|||||
244 ATCTGGCAATGAACCTTAGCGCGAGGTAAGAATCAGCCGACAAA 293

48 sIleThrSerTrpMetGluProIleValLysPheProGlyAlaValAspG 65
|||||
294 AGTA...TCATGG.....GATCGGAACCTG 316

65 lValaThrTyrIleLeuValMetValAspProAspAlaProSerArgAla 81
|||||
317 GAGCCTTATATACGCTCGTTATGACTGATCGGAGCAGCACCATCTCGAAA 366

82 GluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysG 98
|||||
367 AACCCCGTATTACAGAGAGTGGCACCATTGTTGATAATTAATTTCTGG 416

98 yAlaAspLeuLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyr. 114
|||||
417 ACAAAATGTTAGCAGTGGCACAGTG.....TTATCTGATTATT 454

115 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPheP 131
|||||
455 GGATCA...GGTCCACGAAAAAGGCACAGGACTTCATCGTTATGTTCTTT 501

131 eValTyrLeuGlnGluLysValIleSerLeuLeuProLysGluAsnL 148
|||||
502 GGTTTATAAACACCTGGAAGTATCAGGAT.....ACTCACATGGCG 545

148 ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu 164
|||||
546 GAAATCGCCGNAATTTCAAAGTTATGATTTTGCACAAACAAACATCTTG 595

165 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsn 177
|||||
596 GGAATCCAGTTGCGGAACACTTCTTCCAGGCTAAACAT 634
```

```
seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-1
seq_documentation_block:
; Sequence 1, Application US/09060726A
; Patent No. 6225530
; GENERAL INFORMATION:
; APPLICANT: Weigel, Detlef
; APPLICANT: Salk Institute
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; TITLE OF INVENTION: MODIFIED PLANTS HAVING MODULATED FLOWER
; FILE REFERENCE: SALKINS.026A
; CURRENT APPLICATION NUMBER: US/09/060.726A
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: CDS
; FEATURE:
; LOCATION: (63)...(590)
US-09-060-726A-1
```

```
alignment_scores:
  Quality: 164.00      Length: 113
  Ratio: 2.412        Gaps: 4
Percent Similarity: 60.177 Percent Identity: 36.283
```

```
alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-060-726A-1 ..
```

Align seg 1/1 to: US-09-060-726A-1 from: 1 to: 855

```
68 TyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProAr 84
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 TATACCTTGGTTATGGTGATCCAGATGTTCCAAAGCTCTAGCAACCTCA 304
84 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 101
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
305 CTTCCGAGAATATCTCCATTTGGTGGTGAATGATATCCCTGCTACA... 350
101 euLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaPro 117
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
351 .....ACTGGAACAACCTTTGGCAATGAGATTGTGTACGAAATCCA 395
118 SerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
396 AGTCCCACTGCA.....GGAATTCATCGTGTGCTGTTTATATTGTTTCG 439
134 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 ACAGCTTGGCAGGCAACAACAGTGTATGCACCA..... 470
151 lySerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 163
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
471 ..GGGTGGCCGACAGACTTCAACACTCGCGAGTTTCTGAGATCTACAAT 518
164 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 176
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 CTGCGCCTTCCCGTGGCGCAGTTTCTACAATTGTCCAG 557
```

```
seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-09-060-726A-3
```

```
seq_documentation_block:
; Sequence 3, Application US/09060726A
; Patent No. 6225530
; GENERAL INFORMATION:
; APPLICANT: Weigel, Detlef
; APPLICANT: Salk Institute
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
```

```
; TITLE OF INVENTION: MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT
; FILE REFERENCE: SALKINS.026A
; CURRENT APPLICATION NUMBER: US/09/060.726A
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-060-726A-3
```

```
alignment_scores:
  Quality: 164.00      Length: 113
  Ratio: 2.412        Gaps: 4
Percent Similarity: 60.177 Percent Identity: 36.283
```

```
alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-060-726A-3/rev ..
```

Align seg 1/1 to reverse of: US-09-060-726A-3 from: 1 to: 856

```
68 TyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProAr 84
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 TATACCTTGGTTATGGTGATCCAGATGTTCCAAAGCTCTAGCAACCTCA 553
84 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 101
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
552 CTTCCGAGAATATCTCCATTTGGTGGTGAATGATATCCCTGCTACA.... 507
101 euLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaPro 117
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 .....ACTGGAACAACCTTTGGCAATGAGATTGTGTACGAAATCCA 462
118 SerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
461 AGTCCCACTGCA.....GGAATTCATCGTGTGCTGTTTATATTGTTTCG 418
134 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 ACAGCTTGGCAGGCAACAACAGTGTATGCACCA..... 387
151 lySerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 163
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
386 ..GGGTGGCCGACAGACTTCAACACTCGCGAGTTTCTGAGATCTACAAT 339
164 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 176
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 CTGCGCCTTCCCGTGGCGCAGTTTCTACAATTGTCCAG 300
```

```
seq_name: /cgnl_7/ptodata/1/ina/6B_COMB.seq:US-08-998-416-1021
```

```
seq_documentation_block:
; Sequence 1021, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
```

```
; APPLICANT: Philippesen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
```



```
seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-47
seq_documentation_block:
; Sequence 47, Application US/08313185
; Patent No. 5851763
```

99 AlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSer..... 112
 2154203GGCAGCACCGCCCGATGTTGAGACTCCCGGTGGCGGAATCAGCCTGCCGAA 2154154

```

; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356-0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-313-185-47

```

```

alignment_scores:
  Quality: 87.50      Length: 108
  Ratio: 1.620       Gaps: 7
Percent Similarity: 50.000 Percent Identity: 31.481

```

```

alignment_block:
US-09-215-435-225_copy_23_227 x US-08-313-185-47 ..
Align seg 1/1 to: US-08-313-185-47 from: 1 to: 700

42 CysAsnAspTyrArgGlnLysIleThrSertirpMetGluProIleVal... 57
|||||...::: ||| ||| |||...
321 TGCATTCAGTCGAAGGGGCCACG...TGGCGCCACCGTTCACGTG 367
58 .LysPheProGlyAlaValAspGlyAlaThrTyrIleValMetVala 74
::: |||...::: ||| ||| |||...
368 GTCGTCGCCCGCGCGGAGGCAACTGCA.....CTCGTCGTCGATG 408
74 sProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 90
|||||...::: ||| ||| |||...
409 ACCCGACGCGGTGCGC.....GGACTGACGTGCAC 440
91 TrpLeuValThrAspIle.....LysGlyAlaAsp..... 100
|||||...::: ||| ||| |||...
441 TGGATGTCGACCGGAATCGCCCTGCTGGCAGCAGCGCGGATGGTCA 490
101 .....LeuLysLysGlyLysIleGlnGlyG 109
::: ||| ||| ||| |||

```

```

491 GACTCTGCTGCTGGTGGGCACACGCGTCCGAATTCGTGGTGGTGGCAAGGA. 539
109 InGluLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPhe 125
||| ::||| ||||||| ::|||
540 .....TACTTCGGTCCATGCGCGCGCGGCGGACCGGGACA 575
126 HisArgTyrGlnPhePheValTyr 133
|||||...::: ||| ||| |||
576 CACCACTACCGGTTTACCCTCTAC 599
seq_name: /cgnl_7/ptodata/1/ina/5b_COMB.seq:US-08-459-499-11

seq_documentation_block:
; Sequence 11, Application US/08459499
; Patent No. 5871912
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart T.
; APPLICANT: Young, Douglas B.
; APPLICANT: Zhang, Ying
; TITLE OF INVENTION: Nucleic Acid Probes, Sequences, and Methods
; TITLE OF INVENTION: for Detecting Mycobacterium Tuberculosis Resistant to Isoni
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,499
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/875,940
; FILING DATE: 30-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/929,206
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,655
; FILING DATE: 11-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0110-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-459-499-11

```

```

alignment_scores:
  Quality: 87.50      Length: 108
  Ratio: 1.620       Gaps: 7
Percent Similarity: 50.000 Percent Identity: 31.481
alignment_block:

```



```

; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 47:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH: 700 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     US-09-082-614A-47

alignment_scores:
    Quality: 87.50      Length: 108
    Ratio: 1.620      Gaps: 7
    Percent Similarity: 50.000      Percent Identity: 31.481

alignment_block:
    US-09-215-435-225_COPY_23_227 x US-09-082-614A-47      ..
    Align seg 1/1 to: US-09-082-614A-47 from: 1 to: 700

42 CysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleVal... 57
|||||: : : : ||| ||| |||: : : :
321 TGC AATTCAGTCTCAAGGGGGCCCAACG...TGGCCGCCACCGTTCACGTG 367
|||||: : : : ||| ||| |||: : : :
58 LysPheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValA 74
|||||: : : : ||| ||| |||: : : :
368 GTCTGTCGCCCGCGGCGAGCACTGGCA.....CTCGTCGTCGATG 408

74 spProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 90
|||||: : : : ||| ||| |||: : : :
409 ACCCGCGACCGGTCGGC.....GGACTGTACTCGTCGAC 440
|||||: : : : ||| ||| |||: : : :

```

```

91 TrpLeuValThrAspIle.....LysGlyAlaAasp..... 100
|||:::||||| ||| ::| |||||
441 TGGATCGTGACCGGAATCGCCCTGGGTCTGGCAGCACGCCGGATGGTCA 490
101 .....LeuLYSLysGlyLysIleGlnGlyG 109
:::|::|::| |::|::|
491 GACTCTCTGCTGGTGGGCACACGTCGCCGAATCTGGTGGTCGGCAGA. 539
||| ::||| |||||::| :|||
109 lnGluLeuSerAlaTyrglnAlaProserProAlaHisSerGlyphe 125
|||||::|::|::|::|::|::|::|
540 .....TACTTCGGTCCATCGCGCGCGCGGGCACCGGGACA 575
126 HisArgTyrglnPhePheValtyr 133
|||||::|::|::|::|::|::|
576 CACCACCTACCGGTTTACCCTCTAC 599

seq_name: /cgnl_7/ptodata/1/ina/5B_COMB.seq;US-08-459-499-8
seq_documentation_block:
; Sequence 8, Application US/08459499
; Patent No. 5871912
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart T.
; APPLICANT: Young, Douglas B.
; APPLICANT: Zhang, Ying
; TITLE OF INVENTION: Nucleic Acid Probes, Sequences, and Methods
; TITLE OF INVENTION: for Detecting Mycobacterium Tuberculosis Resistant to Isoni
; TITLE OF INVENTION: Amended)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,499
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/875,940
; FILING DATE: 30-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/929,206
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,655
; FILING DATE: 11-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0110-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4794 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-459-499-8

alignment_scores:
    Quality: 87.50      Length: 108
    Ratio: 1.620       Gaps: 7
    Percent Similarity: 50.000    Percent Identity: 31.481

alignment_block:
US-09-215-435-225_copy_23_227 x US-08-459-499-8  ..
Align seg 1/1 to: US-08-459-499-8 from: 1 to: 4794

42 CysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleVal.. 57
321 TGCATTCAGCTGCAAGGGGGCAACG...TGCGCGCCACCGTTGACGTG 367

58 LysPheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValA 74
368 GTCGTGCGCGCGGGGAGCAAGCACTGCA.....CTCGTCGTCGATG 408

74 spProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 90
409 ACCCGACGCGGTCGCGC.....GGACTGTACGTGCAC 440

91 TrpLeuValThrAspIle.....LysGlyAlaAsp..... 100
441 TGGATCGTGACCGGAATCGCCCTGCTGTCGACGACGCGGATGGTCA 490

101 .....LeuLysLysGlyLysIleGlnGly 109
491 GACTCTGCTGGTGGCACAGCGTGCAGATTCGTGTGTCGCGCAAGGA. 539

109 InGluLeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPhe 125
540 .....TACTTCGTGTCATGCGCGCGCGGACCGGGACA 575

126 HisArgTyrGlnPhePheValTyr 133
576 CACCACACTACGGTTTACCTCTTAC 599

seq_name: /cgn1_7/ptodata/1/ina/5B_COMB.seq:US-08-313-185-45

```

```

seq_documentation_block:
; Sequence 45, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/313,185
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-313-185-45

alignment_scores:
    Quality: 87.50      Length: 108
    Ratio: 1.620       Gaps: 7
    Percent Similarity: 50.000    Percent Identity: 31.481

alignment_block:
US-09-215-435-225_COPY_23_227 x US-08-313-185-45  ..
Align seg 1/1 to: US-08-313-185-45 from: 1 to: 4795

42 CysAsnAsnTyrArgGlnLysIleThrSerTrpMetGluProIleVal.. 57
321 TGCATTCAGCTGCAAGGGGGCAACG...TGCGCGCCACCGTTGACGTG 367

58 LysPheProGlyAlaValAspGlyAlaThrTyrIleLeuValMetValA 74
368 GTCGTGCGCGCGGGGAGCAAGCACTGCA.....CTCGTCGTCGATG 408

74 spProAspAlaProSerArgAlaGluProArgGlnArgPheTrpArgHis 90
409 ACCCGACGCGGTCGCGC.....GGACTGTACGTGCAC 440

91 TrpLeuValThrAspIle.....LysGlyAlaAsp..... 100

```

```
441 TGGATCGTGACCGGAATCGCCCTGGCTCTGGCAGCACGCGCGGATGGTCA 490
101 .....LeuLysLysGlyLysIleGlnGlyG 109
      ::::|||||
491 GACTCCTGCTGGTGGGCACAGCGTGCCTGCTGGTGGTGGTGGCAAGGA. 539
109 InGluLeuSerAlaTyrGlnAlaProSerProAlaHisSerGlyPhe 125
      ||| ::::|||||
540 .....TACTTCGGTCCATGCCCGCGCGGGCACCGGGACA 575
126 HisArgTyrGlnPheValTyr 133
      |||:||||:||||| ::::|||||
576 CACCACCTACCGGTTTACCCTCTAC 599
```

THIS PAGE BLANK (USPTO)


```

|||||
231 TCCTGGATGAGCGGATGATCAAGTTCCCGGGGCGGTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCCTGGTGGATGGATCCAGATGCCCTACAGACGAGACCCA 330
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
331 GACAGAGATCTGGAGACATTTGGCTGTTAACAACATACAGACAGAATCACC 380
101 LeuLysLysGlyLysIleGlnGlnGlnLeuLeuSerAlaTyrGlnAlaPr 117
381 CTGAAGAAAGGAGATTTCAAGGGCCAGGAGTTATCAGCCTTACAGGCTCC 430
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
431 CTCGCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTCTTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
481 TTCAGGAAGGAGGATCTCTCTCCATCCCAAGGAAACAAACTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
531 GGCTCTTGGAAATGGACAGATTCTGAACCGTTCCACCTGGCGCAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
581 TGAAGCAAGCACCAGTTTCATGACCCAGAACTACCAAGGACTCACCACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
631 TCCAGGCTCCAGAGAAAGGCGCCAGCGCCCAAGCACAACAAACACGGCG 680
201 GluIleAlaAlaCys 205
681 GAGATAGCTGCCTGC 695

```

seq_name: /cgnl_7/ptodata/1/pna/US094_COMB.seq:us-09-471-276-7

```

seq_documentation_block:
; Sequence 7, Application US/09471276
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: GENSET.025CP1
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pn
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15...695
; NAME/KEY: sig_peptide
; LOCATION: 15...80
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMVMVVG/DE
US-09-471-276-7

```

```

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-471-276-7 ..
Align seg 1/1 to: US-09-471-276-7 from: 1 to: 826
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
81 GACGAGGATGACACAGCCGCGTGTGCCATGAGGCCCTCCTGGACGAGGA 130
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
131 CACCCTCTTTTCCAGGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
181 TTGGCTGCAAGGTTGTTCTCTGATTGTAACAACATACAGACAGAATCACC 230
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
231 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGGTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
281 CTATATCCTGGTGGATGGATCCAGATGCCCTAGCAGACGAGACCAACA 330
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
331 GACAGAGATCTGGAGACATTTGGCTGTTAACAACATATCAAGGGCGCGCAC 380
101 LeuLysLysGlyLysIleGlnGlnGlnLeuLeuSerAlaTyrGlnAlaPr 117
381 CTGAAGAAAGGAGGATTTCAAGGGCCAGGAGTTATCAGCCTTACAGGCTCC 430
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
431 CTCGCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTCTTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
481 TTCAGGAAGGAAAGGTCATCTCTCTCCATCCCAAGGAAACAAACTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
531 GGCTCTTGGAAATGGACAGATTCTGAACCGTTTCCACCTGGCGCAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
581 TGAAGCAAGCACCAGTTTCATGACCCAGAACTACCAAGGACTCACCACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
631 TCCAGGCTCCAGAGAAAGGCGCCAGCGCCCAAGCACAACAAACACGGCG 680
201 GluIleAlaAlaCys 205
681 GAGATAGCTGCCTGC 695

```

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-834-366-7

```

seq_documentation_block:
; Sequence 7, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

```

```

; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834.366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
US-09-834-366-7

```

```

alignment_scores:
  Quality: 1115.00      Length: 205
  Ratio: 5.439          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-215-435-225_COPY_23_227 x US-09-834-366-7 ..

Align seg 1/1 to: US-09-834-366-7 from: 1 to: 826

```

1  AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
81  GAGGAGGATGAGAACAGCCGCTGTCATGAGGCGCTCTCTGGACGAGGA 130
17  pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
131  CACCCCTCTTTGCCAGGGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34  leGlyCysLysValValProAspCysAsnTyrArgGlnLysIleThr 50
181  TTGGCTGCAAGGTTGTCTCTGATTGTAACAACCTACACAGACAGATCACC 230
51  SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlath 67
231  TCTGTGATGAGCCGATAGTCAAGTTCCCGGGGCGGTGGACGGCGAAC 280
67  rTyrlleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
281  CTATATCTGGTGATGGTGATCCACATGCCCTAGCAGACAGAACCCA 330
84  rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
331  GACAGAGATTCTGGACACATCTGGCTGCTAACAGATATCAAGGGCGCCGAC 380
101  LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
381  CTGAAGAAAGGGAAGATTCAAGGCGGAGGAGTTATCAGCTACCGAGCTCC 430
117  oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
431  CTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTTGTCTATC 480
134  euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
481  TTCAGGAAGAAAGGTCATCTCTCTCTCCCAAGGAAACAAACAACTCGA 530
151  GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
531  GGCTCTTGGAAAATGACAGATTTCTGAAACCGGTTTCCACCTGGGCGAAC 580
167  oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184

```

```

|||||
581  TGAAGCAAGCACCACCTTTCATGACCCAGAACTACCAAGACTACCAACCC 630
|||||
184  euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631  TCCAGGCTCCAGAGAAAGGCGCAGCGAGCCCAAGCAACCAAAACCAAGCG 680
|||||
201  GluileAlaAlaCys 205
|||||
681  GAGATAGCTGCCCTGC 695

seq_name: /cgnl_7/ptodata/1/pna/US6008_COMB.seq:US-60-081-563-70

seq_documentation_block:
; Sequence 70, Application US/60081563
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SECRETED PROTEINS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/081.563
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.027PR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Prostate
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
; FEATURE:
; NAME/KEY: poly_a_signal
; LOCATION: 795..800
; FEATURE:
; NAME/KEY: poly_a
; LOCATION: 814..826
; IDENTIFICATION METHOD: blastn
; FEATURE:
; NAME/KEY: est
; LOCATION: 426..664
; IDENTIFICATION METHOD: blastn

```

```
OTHER INFORMATION: identity 99
OTHER INFORMATION: region 99..337
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 327..435
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 1..109
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 716..777
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 100
OTHER INFORMATION: region 391..452
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 660..720
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 334..394
OTHER INFORMATION: id W37255
FEATURE:
NAME/KEY: est
LOCATION: 506..630
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 90
OTHER INFORMATION: region 327..451
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 323..435
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 94
OTHER INFORMATION: region 143..255
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 180..280
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 2..102
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 464..528
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 284..348
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 279..321
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 100
OTHER INFORMATION: region 100..142
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 426..466
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 97
OTHER INFORMATION: region 245..285
OTHER INFORMATION: id W32197
FEATURE:
NAME/KEY: est
LOCATION: 358..674
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 141..457
OTHER INFORMATION: id AA192427
```

```
FEATURE:
NAME/KEY: est
LOCATION: 670..779
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 37..146
OTHER INFORMATION: id AA192427
FEATURE:
NAME/KEY: est
LOCATION: 485..602
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 224..341
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 601..709
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 91
OTHER INFORMATION: region 115..223
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 704..779
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 46..121
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 392..436
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 97
OTHER INFORMATION: region 393..437
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 429..465
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 94
OTHER INFORMATION: region 363..399
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 464..493
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 334..363
OTHER INFORMATION: id W33189
FEATURE:
NAME/KEY: est
LOCATION: 343..730
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 98
OTHER INFORMATION: region 8..395
OTHER INFORMATION: id HSPD04880
FEATURE:
NAME/KEY: est
LOCATION: 507..663
IDENTIFICATION METHOD: blastn
OTHER INFORMATION: identity 96
OTHER INFORMATION: region 264..420
OTHER INFORMATION: id AA192426
US-60-081-563-70
```

```
alignment_scores:
Quality: 1115.00      Length: 205
Ratio: 5.439         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-081-563-70
```



```
Align seg 1/1 to: US-60-081-563-70 from: 1 to: 826
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
81 GACGAGGATGAGAACAGCCGCTGCTGCCATGAGGCCCTCTCTGGACGAGGA 130
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
131 CACCCCTCTTTTGGCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
181 TTGGCTGCAAGGTGTCTCTGATTGTAACAACCTACAGACAGAGATCACC 230
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
231 TCCTGGATGGAGCGGATGTAAGTTCCCGGGCGCTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
281 CTATATCCTGGTGGTGGATCCAGATGCCCTACGAGCAGAACCCCA 330
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
331 GACAGAGATTCTGGAGACATTTGGCTGGTAACAGATATCAAGGGCGCGAC 380
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
381 CTGAAGAAAGGAAAGTCTCTCTCTCCCAAGGAAACAAACCTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyLupr 167
531 GGCTCTTGGAAATGACAGATTCTGAACCGTTTCCACCTGGGCGAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
581 TGAAGCAAGCACCAGTTTATGACCCAGAACTACAGGACTACCAACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
631 TCCAGGCTCCAGAGAAAGGCGCAGCGAGCCCAAGCACAACAAACAGGCG 680
201 GluIleAlaAlaCys 205
681 GAGATAGCTGCCTGC 695
```

seq_name: /cgml_7/ptodata/1/pna/US6014_COMB.seq:US-60-147-499-7

```
seq_documentation_block:
; Sequence 7, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147.499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLLGLMMVVVG/DE
US-60-147-499-7

alignment_scores:
  Quality: 1115.00      Length: 205
  Ratio: 5.439          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-147-499-7 ..
Align seg 1/1 to: US-60-147-499-7 from: 1 to: 826
```

```
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
81 GACGAGGATGAGAACAGCCGCTGCTGCCATGAGGCCCTCTCTGGACGAGGA 130
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
131 CACCCCTCTTTTGGCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
181 TTGGCTGCAAGGTGTCTCTGATTGTAACAACCTACAGACAGAGATCACC 230
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
231 TCCTGGATGGAGCGGATGTAAGTTCCCGGGCGCTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
281 CTATATCCTGGTGGTGGATCCAGATGCCCTACGAGCAGAACCCCA 330
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
331 GACAGAGATTCTGGAGACATTTGGCTGGTAACAGATATCAAGGGCGCGAC 380
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
381 CTGAAGAAAGGAAAGTCTCTCTCTCCCAAGGAAACAAACCTCGA 430
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
431 CTCCCAACCGGCACACAGTGGCTTCCATCGCTACCACTTCTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
481 TTCAGGAAGGAAAGGTCATCTCTCTCTCCCAAGGAAACAAACCTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyLupr 167
531 GGCTCTTGGAAATGACAGATTCTGAACCGTTTCCACCTGGGCGAACC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
581 TGAAGCAAGCACCAGTTTATGACCCAGAACTACAGGACTACCAACCC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
631 TCCAGGCTCCAGAGAAAGGCGCAGCGAGCCCAAGCACAACAAACAGGCG 680
201 GluIleAlaAlaCys 205
681 GAGATAGCTGCCTGC 695
```

seq_name: /cgnl_7/ptodata/1/pna/US6016_COMB.seq:US-60-169-629-7

```
seq_documentation_block:
; Sequence 7, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLLGLMMVVVG/DE
US-60-169-629-7
```

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_23_227 x US-60-169-629-7 ..

Align seg 1/1 to: US-60-169-629-7 from: 1 to: 826

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GACGAGGATGAGACAGCCCGTGTCCCATGAGGCCCTCTGGACGAGGA 130
|||||
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCCTCTTTGCCAGGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
|||||
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTTGTTCTGTGATTGTAACACTACAGACAGAGATCACC 230
|||||
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGGCGGTGGACGGCGAAC 280
|||||
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCTGGTGGTGGATCCAGATGCCCTTAGCAGACAGAACCCCA 330
|||||
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATTCGTGAGACATTTGCTGCTGAACAGATATCAAGGGCGCGCAC 380
|||||
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrClnAlaPr 117
|||||
381 CTGAAGAAGGGGAAGATTCAGGGCCAGGAGTTATCAGCCTACCGAGCTCC 430
|||||
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
431 CTCCCCACCGGCACACAGTGGCTTCCATCGCTACCGATCTTGTCTATC 480
|||||

134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAAGGAAGGTGATCTCTCTCCCTCCCAAGAGAAAACAACTCGA 530
|||||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
531 GGCTCTTGGAATGACAGATTTCTGAACCGTTTCCACCTGGGGAACC 580
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAAGCACCCAGTTTCATGACCCAGAACTACCAAGGACTACCAACCC 630
|||||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCCAGCGAGGCCCAAGCACAAAAAACCCAGGG 680
|||||
201 GluIleAlaAlaCys 205
|||||
681 GAGATAGTCGCTGC 695
|||||

seq_name: /cgnl_7/ptodata/1/pna/US6018_COMB.seq:US-60-187-470-7

```
seq_documentation_block:
; Sequence 7, Application US/60187470
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAs for Secreted Proteins
; FILE REFERENCE: 78.US2.PRO
; CURRENT APPLICATION NUMBER: US/60/187,470
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLLGLMMVVVG/DE
US-60-187-470-7
```

alignment_scores:
Quality: 1115.00 Length: 205
Ratio: 5.439 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-215-435-225_COPY_23_227 x US-60-187-470-7 ..

Align seg 1/1 to: US-60-187-470-7 from: 1 to: 826

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GACGAGGATGAGACAGCCCGTGTGCCCATGAGGCCCTCTGGACGAGGA 130
|||||
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCCTCTTTGCCAGGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
|||||
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTTGTTCTGTGATTGTACAACTACAGACAGAGATCACC 230
|||||
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||

```

|||||
231 TCCTGGATGGAGCCGATAGTCAAGTTCCGGGGCGCGTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCCTGGTGGTGGATGCCAGATGCCCTAGCAGAGCAGAACCCA 330
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATTCGGAGACATTCGGCTGGTAAACAGATATCAAGGGCGCGAC 380
101 LeuLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaLpr 117
|||||
381 CTGAAGAAAGGAAAGATTCAGGGCAGGAGTTATCAGCCTACCGGCTCC 430
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
431 CTCGCCACGGGCACACAGTGGCTTCATCGCTACCATGTTCTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAAGAAAGGTCATCTCTCTCCCAAGGAAACAAACCTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyLpr 167
|||||
531 GGCCTCTTGGAAATGGACAGATTTCTGAACCGTTTCCACCTGGGCGAAC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAACACCCAGTTCATGACCCAGAACTACCAAGGCGCGAC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCGCAGGAGCCCAAGCACAACCAAGGCG 680
201 GluIleAlaAlaCys 205
|||||
681 GAGATAGTGCCTGCG 695

```

seq_name: /cgnl_7/ptodata/1/pna/US6019_COMB.seq:US-60-197-873-7

```

seq_documentation_block:
; Sequence 7, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pmp
; SEQ ID NO 7
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig.peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVVVG/DE
US-60-197-873-7

```

```

alignment_scores:
Quality: 1115.00
Ratio: 5.439
Length: 205
Gaps: 0

```

```

Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-197-873-7
Align seg 1/1 to: US-60-197-873-7 from: 1 to: 826
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
81 GACGAGGATGAGAACAGCCGCTGTGCCCATGAGGCCCTCTCTGGACGAGGA 130
17 PheLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCCCTCTTTTCCAGGGCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 180
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTGTTCTCTGATTGTAAACAATACACAGACAGAGATCAC 230
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGCTGGACGGCGCAAC 280
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCCTGGTGGTGGATGCCAGATGCCCTAGCAGAGCAGAACCCA 330
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATTCGGACACATTCGGCTGGTAAACAGATATCAAGGGCGCGAC 380
101 LeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaLpr 117
|||||
381 CTGAAGAAAGGGAAGATTCAGGGCCAGGAGTTATCAGCCTACCGGCTCC 430
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
431 CTCGCCACGGGCACACAGTGGCTTCATCGCTACCATGTTCTTGTCTATC 480
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAAGAAAGGTCATCTCTCTCCCAAGGAAACAAACCTCGA 530
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyLpr 167
|||||
531 GGCCTCTTGGAAATGGACAGATTTCTGAACCGTTTCCACCTGGGCGAAC 580
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGCAACACCCAGTTCATGACCCAGAACTACCAAGGCGCGAC 630
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCAGAGAAAGGCGCAGGAGCCCAAGCACAACCAAGGCG 680
201 GluIleAlaAlaCys 205
|||||
681 GAGATAGTGCCTGCG 695

```

```

seq_name: /cgnl_7/ptodata/1/pna/US6022_COMB.seq:US-60-224-006-7
seq_documentation_block:
; Sequence 7, Application US/60224006
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: ESTS encoding human proteins
; FILE REFERENCE: 95.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/224,006
; CURRENT FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent.pmp
; SEQ ID NO 7

```

```

; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..695
; NAME/KEY: sig_peptide
; LOCATION: 15..80
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.5
; OTHER INFORMATION: seq AALLGLMMVTG/DE
; US-60-224-006-7

```

```

alignment_scores:
  Quality: 1115.00      Length: 205
  Ratio: 5.439          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-215-435-225_COPY_23_227 x US-60-224-006-7 ..

Align seg 1/1 to: US-60-224-006-7 from: 1 to: 826

```

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuAspGluAs 17
|||||
81 GACGAGGATGAGAACAGCCCGTGTGCCATGAGGCCCTCTGGACGAGGA 130
|||||
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
131 CACCTCTTTTCCAGGCGCTTGAGTTTCTACCCAGAGTTGGGGAACA 180
|||||
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
181 TTGGCTGCAAGGTTGTTCTCTGATTGTAACTACAGACAGAGATCACC 230
|||||
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
231 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGCTGGACGGCGAAC 280
|||||
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
281 CTATATCTGTGTGATGGTGGATCCAGATGCCCTAGCAGACAGAACCCA 330
|||||
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
331 GACAGAGATTCGGAGACATTTGGCTGGTACAGATATCAAGGGCGCCGAC 380
|||||
101 LeuLysLysGlyLysIleGlnGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
381 CTGAAGAAAGGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACAGGCTCC 430
|||||
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
431 CTCGCCACCGGCACACAGTGGCTTCATCGCTACCCAGTTCTTTGTCTATC 480
|||||
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
481 TTCAGGAGGAAGGTCATCTCTCTCTCCCAAGGAAACAAACTCGA 530
|||||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
531 GGCTCTTGGAAATGGACAGATTTCTGAACCGTTTCCACCTGGGGGAACC 580
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
581 TGAAGGCAACACCCAGTTCATGACCCAGAACTACCCAGGACTCACCACCC 630
|||||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
631 TCCAGGCTCCACAGAAAGGGCCAGGAGCCCAAGCACAACCAACGAGCG 680
|||||
201 GluIleAlaAlaCys 205

```

```

|||||
681 GAGATAGCTGCCTGC 695

```

seq_name: /cgnl_7/ptodata/1/pna/US6027_COMB.seq:US-60-278-258-12888

seq_documentation_block:

```

; Sequence 12888, Application US/60278258
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0010-1 P
; CURRENT APPLICATION NUMBER: US/60/278, 258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 12888
; LENGTH: 904
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 204392.4
; US-60-278-258-12888

```

```

alignment_scores:
  Quality: 1115.00      Length: 205
  Ratio: 5.439          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-09-215-435-225_COPY_23_227 x US-60-278-258-12888 ..

Align seg 1/1 to: US-60-278-258-12888 from: 1 to: 904

```

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
155 GACGAGGATGAGAACAGCCCGTGTGCCATGAGGCCCTCTTGGACGAGGA 204
|||||
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
205 CACCTCTTTTGGCAGGGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 254
|||||
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
255 TTGGCTGCAAGGTTGTTCTGTATTGTAACAACCTACAGACAGAGATCACC 304
|||||
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
305 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCGCTGGACGGCGCAAC 354
|||||
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
355 CTATATCTGTGTATGGTGGATCCAGATGCCCTAGCAGACAGAACCCA 404
|||||
84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
405 GACAGAGATTCCTGAGACATTTGGCTGGTAAACAGATATCAAGGGCGCCGAC 454
|||||
101 LeuLysLysGlyLysIleGlnGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
455 CTGAAGAAAGGGAAGATTTCAGGGCCAGGAGTTTATCACCTACCCAGGCTCC 504
|||||
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
505 CTCGCCACCGGCACACAGTGGCTTCCATCTGCTACCCAGTTCTTTGTCTATC 554
|||||
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||

```



```

Percent Similarity: 99.512 Percent Identity: 99.512
alignment_block:
US-09-215-435-225_COPY_23_227 x PCT-US99-29950-56
Align seg 1/1 to: PCT-US99-29950-56 from: 1 to: 1028

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
219 GACGAGGATGAGAACACGCCGTGTGCCCATGAGGCCCTCTGGACGAGGA 268

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
269 CACCTCTCTTTGCCAGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 318

34 LeGlyCysLysValValProAspCysAsnTyrArgGlnLysIleThr 50
|||||
319 TTGGCTGCAAGTTGTTCTGTGATTGTAACTACACAGAGATCACC 368

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
369 TCCTGGATGAGCGGATAGTCAAGTTCCCGGGCGCTGGACGGCGAAC 418

67 rTyrlleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
419 CTATATCTGTGTGATGGTGGTCAAGTTCCCGGGCGCTGGACGGCGAAC 468

84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
469 GACAGAGATTCGAGACATTTGGCTGGTAAACAGATATCAAGGGCGCGAC 518

101 LeuLysLysGlyLysIleGlnGlyGlnLeuLeuSerAlaTyrGlnAlaPr 117
|||||
519 CTGAAGRAAGGAGAGATTCAAGGGCGGAGGTTATCAGCCTACCGGCTCC 568

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
569 CTCCACCGGACACAGTGGCTCCATCGCTACAGTCTTCTGTCTATC 618

134 euGlnGluGlyLysValLysSerLeuLeuProLysGluAsnLysThrArg 150
|||||
619 TTCAGGAAGGAAAGTCACTCTCTCTCCCTCCCAAGGAAACAAACATCGA 668

151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
669 GGCTCTTGAAATGGACAGATTCTGAACTTTCCACCTTGGCGGAC 718

167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
719 TCAAGCAAGCAGCCAGTTTCATGACCCAGAACTTACCGAGACTCAGCAACC 768

184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
769 TCCAGGCTCCAGAGAGGGCGGACGAGCCCAAGCAAAACAGGCG 818

201 GluIleAlaAlaCys 205
|||||
819 GAGATAGCTGCCTGC 833

seq_name: /cgn1_7/ptodata/1/pna/US095D_COMB.seq:US-09-591-316-56

seq_documentation_block:
; Sequence 56, Application US/09591316
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P203591
; CURRENT APPLICATION NUMBER: US/09/591.316
; CURRENT FILING DATE: 2000-06-09
; EARLIER APPLICATION NUMBER: PCT/US99/29950
; EARLIER FILING DATE: 1999-12-16
; EARLIER APPLICATION NUMBER: 60/113,006
; EARLIER FILING DATE: 1998-12-18

```

```

; EARLIER APPLICATION NUMBER: 60/112,809
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1022)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1026)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1027)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1028)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-591-316-56

alignment_scores:
Quality: 1109.00 Length: 205
Ratio: 5.436 Gaps: 0
Percent Similarity: 99.512 Percent Identity: 99.512

alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-591-316-56
Align seg 1/1 to: US-09-591-316-56 from: 1 to: 1028

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
219 GACGAGGATGAGAACACGCCGTGTGCCCATGAGGCCCTCTGGACGAGGA 268

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
269 CACCTCTCTTTGCCAGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 318

34 LeGlyCysLysValValProAspCysAsnTyrArgGlnLysIleThr 50
|||||
319 TTGGCTGCAAGTTGTTCTGTGATTGTAACTACACAGAGATCACC 368

51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
369 TCCTGGATGAGCGGATAGTCAAGTTCCCGGGCGCTGGACGGCGAAC 418

67 rTyrlleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
419 CTATATCTGTGTGATGGTGGTCAAGTTCCCGGGCGCTGGACGGCGAAC 468

84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
469 GACAGAGATTCGAGACATTTGGCTGGTAAACAGATATCAAGGGCGCGAC 518

101 LeuLysLysGlyLysIleGlnGlyGlnLeuLeuSerAlaTyrGlnAlaPr 117
|||||
519 CTGAAGRAAGGAGAGATTCAAGGGCGGAGGTTATCAGCCTACCGGCTCC 568

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
|||||
569 CTCCACCGGACACAGTGGCTCCATCGCTACAGTCTTCTGTCTATC 618

134 euGlnGluGlyLysValLysSerLeuLeuProLysGluAsnLysThrArg 150
|||||
619 TTCAGGAAGGAAAGTCACTCTCTCTCCCTCCCAAGGAAACAAACATCGA 668

619 TTCAGGAAGGAAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 668

```

151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
669 GGCTCTTGAAAATGACAGATTCTGAACCGTTTCCACCTGGCGGAACC 718
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
719 TGAAGCAAGCACCCAGTTCATGACCCAGAACTACCAAGSACTCACCACCC 768
|||||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
769 TCCAGGCTCCAGAGAAAGGGCCAGGAGCCCAAGCACAAAACACGAGCG 818
|||||
201 GluileAlaAlaCys 205
|||||
819 GAGATAGCTGCCTGC 833

seq_name: /cgnl_7/ptodata/1/pna/US098_COMB.seq:US-09-895-298-56

seq_documentation_block:

; Sequence 56, Application US/09895298

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 47 Human Secreted Proteins

; FILE REFERENCE: P2035P1

; CURRENT APPLICATION NUMBER: US/09/895,298

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: 09/591,16

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: PCT/US99/29950

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: 60/113,006

; PRIOR FILING DATE: 1998-12-18

; PRIOR APPLICATION NUMBER: 60/112,809

; PRIOR FILING DATE: 1998-12-17

; NUMBER OF SEQ ID NOS: 231

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 56

; LENGTH: 1028

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (1022)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (1026)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (1027)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (1028)

; OTHER INFORMATION: n equals a,t,g, or c

; US-09-895-298-56

alignment_scores:

Quality: 1109.00

Ratio: 5.436

Percent Similarity: 99.512

Length: 205

Gaps: 0

Percent Identity: 99.512

US-09-215-435-225_COPY_23_227 x US-09-895-298-56

Align seg 1/1 to: US-09-895-298-56 from: 1 to: 1028

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17

|||||

219 GAGCAGGATGAGAACAGCCGCTGCGCCATGAGGCCCTCTTGACCGAGGA 268

|||||

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34

|||||

269 CACCCCTTTTGGCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 318

|||||

34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
319 TTGGCTTGCAAGGTGTTCTCTGATTGTAACAACCTACAGACAGAAGATCACC 368
|||||
51 SerTrpMetCluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
369 TCTGTGATGAGCGGATAGTCAAGTTCCCGGGGGCGGTGACGGCGCAAC 418
|||||
67 rTyrlleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
|||||
419 CTATATCCTGGTGATGGTGATCCAGATCCCTAGCAGAGCAGAACCCA 468
|||||
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
469 GACAGAGATTCTGGACACATTGGCTGGTAACAGATATCAAGGGCGCGGAC 518
|||||
101 LeuLysGlyLysLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
|||||
519 CTGAGRAAGGGAAGATTTCAGGGCCAGGAGTTATCAGCCTTACCAGGCTCC 568
|||||
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
|||||
569 CTCGCCACCGGCACACAGTGGCTTCCATCGCTACCGATTCTTTGTCTATC 618
|||||
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
|||||
619 TTCAGGAAGGAAAGTCATCTCTCTCTCCCAAGGAAACAAACTCGA 668
|||||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
|||||
669 GGCTCTTGGAAAATGGACAGATTCTGAACCGTTTCCACCTGGCGGAACC 718
|||||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
|||||
719 TGAAGCAAGCACCCAGTTCATGACCCAGAACTACCCAGGACTCACCACCC 768
|||||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
|||||
769 TCCAGGCTCCAGAGAAAGGGCCAGGAGCCCAAGCACAAAACACGAGCG 818
|||||
201 GluileAlaAlaCys 205
|||||
819 GAGATAGCTGCCTGC 833

seq_name: /cgnl_7/ptodata/1/pna/US6017_COMB.seq:US-60-172-373-10894

seq_documentation_block:

; Sequence 10894, Application US/60172373

; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonald

; APPLICANT: Lal, Preeti

; APPLICANT: Diep, Dinh

; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using

; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polym

; FILE REFERENCE: GX-0006 P

; CURRENT APPLICATION NUMBER: US/60/172,373

; CURRENT FILING DATE: 1999-12-16

; NUMBER OF SEQ ID NOS: 25,772

; SOFTWARE: PERL Program

; SEQ ID NO 10894

; LENGTH: 904

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No: 204392.1

US-60-172-373-10894

alignment_scores:

Quality: 1108.00

Ratio: 5.431

Length: 205

Gaps: 0

Thu Nov 15 10:52:15 2001

Percent Similarity: 99.512 Percent Identity: 99.512
 alignment_block:
 US-09-215-435-225_COPY_23_227 x US-60-172-373-10894 ...
 Align seg 1/1 to: US-60-172-373-10894 from: 1 to: 904

```

1  AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
  |||
155 GACGAGGATGAGAACACGCCGCTGTGCCATGAGGCCCTCTTGGACGAGGA 204
  |||
17  pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
  |||
205 CACCCCTCTTTGGCCAGGGCCCTTGAAGTTTCTACCCAGAGTTGGGGAACA 254
  |||
34  leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
  |||
255 TTGCTGCAAGGTGTTCTCTGATTGTAACTACACACAGACAGATCACC 304
  |||
51  SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
  |||
305 TCCTGGATGGAGCGGATAGTCAAGTTCCTCCGGGGCCGTGGACGGCGCAC 354
  |||
67  rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
  |||
355 CTATATCCTGGTGATGGTGGATCCAGATGCCCTTAGCGAGACGACACCCA 404
  |||
84  rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
  |||
405 GACAGAGATTCTGGAGACATTGGCGTGGTAAACAGATATCAAGGGCGCGAC 454
  |||
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
  |||
455 CTGAGAAAGGGAAGATTTCAGGGCCAGGAGTTATCAGCCTACCAGGCTCC 504
  |||
117 cSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
  |||
505 CTCCCACCGGCACACAGTGGCTTCCATCGCTACAGTTCTTGTCTATC 554
  |||
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArg 150
  |||
555 TTCAGGAAGGAAAGTCATCTCTCTCTCCCAAGGAAACAAACACTCGA 604
  |||
151 GlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPr 167
  |||
605 GGCTCTTGAAATGGACAGATTCTGAACCGCTTCCACCTGGGCGRACC 654
  |||
167 oGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrL 184
  |||
655 TGAAGCAAGCACCCAGTTTCATGACCCAGAACTACCAGGACTCACCAACC 704
  |||
184 euGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAla 200
  |||
705 TCCAGGCTCCCAAGGAGGGCGGACGAGCCCAAGCAACAAACACGAGCG 754
  |||
201 GluIleAlaAlaCys 205
  |||
755 GAGATAGCTGCCTGC 769
  
```


OM of: US-09-215-435-225_COPY_23_227 to: Pending_Patents_NA_New:* out_format : pfs

Date: Nov 15, 2001 4:11 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cgnl_1/USPTO.spool/US09215435/runat_14112001_141105_2194/app_query.fasta_1.1519
-DB=pending_patents_NA_New -QWTF=fastap -SUFFIX=rnpn
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blosum62 -TRANS=human40 cdi -LIST=A5 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pts -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09215435_@cgnl_1_69 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-215-435-225_COPY_23_227

Query length: 205

Database: Pending_Patents_NA_New:*

Database sequences: 127876

Database length: 122380375

Search time (sec): 564.420000

score_list:

```

Sequence      Strd Orig      zScore      EScore Len      ! Documentation
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-960-253-167 + 205.00 439.49 2.9e-17 1
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-845-849-1 + 164.00 350.79 2.5e-12 856
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-845-849-3 + 164.00 350.79 2.5e-12 856
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-4131 + 158.00 336.41 1.6e-11 5
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-971 + 105.50 220.93 4.3e-05 5
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-1413 + 103.00 221.23 4.2e-05 5
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-3078 + 96.50 198.20 0.0008 5
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-4132 + 92.50 187.32 0.0032 5
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-815-242-6698 + 80.50 156.19 0.1744 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-6710 + 79.00 161.93 0.0835 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-481 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3027 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3028 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3029 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3030 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3031 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3032 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3033 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3034 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3035 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3036 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3037 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3038 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3039 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3040 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3041 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3042 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3043 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3044 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3045 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3046 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3047 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3048 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3049 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3050 + 78.50 112.60 46.71 2
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-3678 + 78.00 154.38 0.2198 1
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-3304 - 77.00 117.17 25.98 1
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2244 - 77.00 117.17 25.98 1
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2245 - 77.00 117.17 25.98 1
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2246 - 77.00 117.17 25.98 1

```

/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2247 - 77.00 117.17 25.98
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2248 - 77.00 117.17 25.98
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2249 - 77.00 117.17 25.98
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2250 - 77.00 117.17 25.98
/cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-897-516-2251 - 77.00 117.17 25.98
seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-960-253-167

seq_documentation_block:
; Sequence 167, Application US/09960253
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960.253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-167

alignment_scores:
Quality: 205.00 Length: 151
Ratio: 2.228 Gaps: 4
Percent Similarity: 60.927 Percent Identity: 33.113

alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-960-253-167 ..
Align seg 1/1 to: US-09-960-253-167 from: 1 to: 1444

30 GluLeuGlyAsnIleGlyCysLysValValProAspCysAsnAsnTyrAr 46
|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
216 GAGCTGGGCAAGTG.....CTGACGCCCCACCCAGGTAAAGTAG 256

46 gGlnLysIleThrSerTrpMetGluProIleValLysPheProGlyAlav 63
|||||
257 ACCCACCACCAITTCGTGG.....GATGGCTTCG 285

63 aLaspGlyIalThrTyrIleLeuValMetValaspProaspAlaProSer 79
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||
286 ATTACGGGAAGCTCTACACCTTGGTCTGACAGACCCGGATGCTCCAGC 335

80 AtgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAsp11 96
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
336 AGAAGGATCCCAATACAGAGTAATGGCATTCCTCTGGTGGTCAACAT 395

96 eLysGlyAlaAspLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerA 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 GAAGGCAATGACATCAGCAGTGGCACAGTC.....CTCTCCG 423

113 lAtyTrGlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGln 129
|||:|||||:|||||:|||||:|||||:|||||:|||||
424 ATTATGGGCTCGGGGCTCCCAAGGCACAGGCTCCACCGCTATGTC 473

130 PhePheValTyrLeuGlnGluGlyLysValIleSer.....Le 142
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
474 TGGCTGGTTTACGACGACGAGCGCGCTAAAGTGTGACGAGCCCATCCT 523

142 uLeuProLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheL 159
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
524 CAGCAACCGATCTGGAGACACCGCTGGCAAAATCAAGTGGGCTCTTCC 573

159 euAsnArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThr 175
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
574 GTAAAAAGTATGAGCTCAGGGGCCCGGCTGGCCTGGCAGCGTGTACCAGGCC 623

```

; APPLICANT: WEIGEL, Detlef
; APPLICANT: KARDAILSKY, Igor
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT
; FILE REFERENCE: SALKINS.026DV1
; CURRENT APPLICATION NUMBER: US/09/845,849
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/060,726
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-845-849-3

```

```

alignment_scores:
    Quality: 164.00    Length: 113
    Ratio: 2.412      Gaps: 4
    Percent Similarity: 60.177    Percent Identity: 36.283

```

```

alignment_block:
US-09-215-435-225_copy_23_227 x US-09-845-849-3/rev ..
Align seg 1/1 to reverse of: US-09-845-849-3 from: 1 to: 856

```

```

68 TyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProAr 84
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 TATACTTTGGTTATGGTGATCCAGATGTTCCAAAGTCTAGCAACCCCA 553
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
84 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 101
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
552 CTTCCGAGAAATATCTCCATTGTTGGTACTGATATCCCTGCTACA.... 507
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
101 euLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPro 117
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
506 .....ACTGGAACAACCTTTGGCAATGAGATTGTGTGTACGAAATCCA 462
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
118 SerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 134
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
461 AGTCCCACTGCA.....GGAATTCATCGTGTGCGTTTATATTGTTTCG 418
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
134 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 151
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
417 ACAGCTTGGCAGGCAACACAGTGTATGCACCA..... 387
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
151 lySerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 163
   :|||: ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
386 ..GGGTGGCGCCAGAACTTCAACACTCGCGAGTTTGTGATGATCTACAAT 339
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
164 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 176
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
338 CTGCGCTTCCCGTGGCGCAGTTTCTACAATTGTTCAG 300
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-60-325-448-4131

```

seq_documentation_block:
; Sequence 4131, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 4131
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-60-325-448-4131

```

```

176 Gln 176
   :|||
624 GAG 626

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-845-849-1

seq_documentation_block:
; Sequence 1, Application US/09845849
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: WEIGEL, Detlef
; APPLICANT: KARDAILSKY, Igor
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT
; FILE REFERENCE: SALKINS.026DV1
; CURRENT APPLICATION NUMBER: US/09/845,849
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/060,726
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-845-849-1

```

```

alignment_scores:
    Quality: 164.00    Length: 113
    Ratio: 2.412      Gaps: 4
    Percent Similarity: 60.177    Percent Identity: 36.283

```

```

alignment_block:
US-09-215-435-225_copy_23_227 x US-09-845-849-1 ..
Align seg 1/1 to: US-09-845-849-1 from: 1 to: 856

```

```

68 TyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProAr 84
   ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 TATACTTTGGTTATGGTGATCCAGATGTTCCAAAGTCTAGCAACCCCTCA 304
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
84 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 101
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
305 CTTCCGAGAAATATCTCCATTGTTGGTACTGATATCCCTGCTACA.... 350
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
101 euLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPro 117
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
351 .....ACTGGAACAACCTTTGGCAATGAGATTGTGTGTACGAAATCCA 395
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
118 SerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 134
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
396 AGTCCCACTGCA.....GGAATTCATCGTGTGCGTTTATATTGTTTCG 439
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
134 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 151
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
440 ACAGCTTGGCAGGCAACACAGTGTATGCACCA..... 470
   ||| ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
151 lySerTrpLysMetAsp.....ArgPheLeuAsnArgPheHis 163
   :|||: ||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
471 ..GGGTGGCGCCAGAACTTCAACACTCGCGAGTTTGTGATGATCTACAAT 518
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
164 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 176
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
519 CTCGCGCTTCCCGTGGCGCAGTTTCTACAATTGTTCAG 557
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-845-849-3

```

seq_documentation_block:
; Sequence 3, Application US/09845849
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES

```

```
alignment_scores:
  Quality: 158.00      Length: 141
  Ratio: 1.859        Gaps: 4
  Percent Similarity: 60.284  Percent Identity: 29.078

alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-325-448-4131 ..
Align seg 1/1 to: US-60-325-448-4131 from: 1 to: 891

54 GluProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLe 70
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
255 AAGCCACGCAATTCAGTCAGGGTGGGACATGAGATCCTTCTCACTCT 304
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

70 uValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgP 87
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
305 GGTCAATGACGGACCCAGATGTGCCAGGGGCTAGTGATCATACCTTAGG 354
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

87 hetrPArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 103
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
355 AGCATCTCCACTGGATCGTCAGTGATATTCCTGGCACCC.....ACA 395

104 GlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaProSerProPr 120
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
396 GATGCTTCTTTTGGAGGGAGGTGGTGAGCTACGAGAGCCCAAGCCC... 443

120 oAlaHisSerGlyPheHisArgTyrGlnPheValTyrLeuGlnGluG 137
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
444 ...AACATTGGCATTCACAGGTTCACTTCGTGCTGTTCACACAGAAGA 489

137 lYsValIleSerLeuLeuProLysGluAsnLysThrArgGlySerTrp 153
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
490 AGCGGAGGCGCATGAACCTCTCC.....TCCACAGGGGACTACTTC 530

154 LysMetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAla 170
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
531 AACACGCGCGCTTCGCCAACGAGACGACCTTGGCTCCCGGTGCTGC 580

170 rThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrLeuGlnAla 187
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
581 CGTCTACTTCAAGCCGACG.....CGGAGACAGCGGACGCGCGCGCT 624

187 roArgGluArgAlaSerGluPro 194
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
625 GACGGAAACAATCAGCCCAACCC 647

seq_name: /cgnl_7/ptodata/1/pna/US60_NEW_COMB.seq:US-60-325-448-971

seq_documentation_block:
; Sequence 971, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325.448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 971
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-971

alignment_scores:
  Quality: 105.50      Length: 113
  Ratio: 1.598        Gaps: 2
  Percent Similarity: 58.407  Percent Identity: 26.549

alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-325-448-971 ..
Align seg 1/1 to: US-60-325-448-971 from: 1 to: 321

71 ValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgPh 87
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
1 GTTATGACAGACCCAGATGTGCCAGGACCAAGTGATCCATATCTAAGGGA 50
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

87 etrPArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 103
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
51 ACACCTACATTG.ATTGTAACGTATATACCTGGACAACACGGATG..... 93

104 GlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaProSerProPr 120
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
94 ...CCTCTTTTGGAGCGGAAATCATAAAGCTATGAGAGCCCAAGCCAG 140
```

```
US-09-215-435-225_COPY_23_227 x US-60-325-448-971 ..
Align seg 1/1 to: US-60-325-448-971 from: 1 to: 535

64 AspGlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerAr 80
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
185 GATCTTTTCTTCACATTGGTGGTTATCACAGACCCAGATGTGCCAGACC 234
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

80 gAlaGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleL 97
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
235 AGTGATCTTATCTAAGGAGCACCTTCATTCATTTAGCGAGGATTGT 284
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

97 ysGlyAlaAspLeuLysLysGlyLysIleGlnGlnGluLeuSerAla 113
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
285 TACTGATATACCTGGGACAAACGGATGCTTTTTCGCGAGGTCATAAGC 334
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

114 TyrGlnAlaProSerProAlaHisSerGlyPheHisArgTyrGlnPh 130
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
335 TATGAGAGTCCAAGCG.....AACATTGGCATCCATAGGTTTCATTTT 378
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

130 ePheValTyrLeuGlnGlnGlyLysValIleSerLeuLeuProLysGlu 147
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
379 TGTGCTCTTCAAGCAGAGCGGACCAAACTGTAATTGTGCCA..... 421
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

147 snLysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHis 163
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
422 ..TCCTTCAGGGACCATTTCAACACCCGCGGTTCGCCGAGGAGATGAT 469
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

164 LeuGlyGluProGluAlaSerThrGlnPheMetThrGln 176
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
470 CTTCGCTTCTCTGCTGCTGCTACTTCAATGCCCCAG 508

seq_name: /cgnl_7/ptodata/1/pna/US60_NEW_COMB.seq:US-60-325-448-1413

seq_documentation_block:
; Sequence 1413, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; APPLICANT: Moughamer, T.G.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325.448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 1413
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-1413

alignment_scores:
  Quality: 103.00      Length: 107
  Ratio: 1.537        Gaps: 4
  Percent Similarity: 62.617  Percent Identity: 30.841

alignment_block:
US-09-215-435-225_COPY_23_227 x US-60-325-448-1413 ..
Align seg 1/1 to: US-60-325-448-1413 from: 1 to: 321

71 ValMetValAspProAspAlaProSerArgAlaGluProArgGlnArgPh 87
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
1 GTTATGACAGACCCAGATGTGCCAGGACCAAGTGATCCATATCTAAGGGA 50
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

87 etrPArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysLys 103
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
51 ACACCTACATTG.ATTGTAACGTATATACCTGGACAACACGGATG..... 93

104 GlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaProSerProPr 120
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
94 ...CCTCTTTTGGAGCGGAAATCATAAAGCTATGAGAGCCCAAGCCAG 140
```

249 ACTGCTGCCAGGAGGGCTGAAAATCCAGCTCT 281

documentation_block:
cc0 Attribution: https://doi.org/10.26434/chemrxiv-2024-8792v

```
seq_name: /cgnl_7/prodata/1/pna/us09_US09_NEW_COMB.seq:US-09-815-242-6698
seq_documentation_block:
seq: Sequence 6698, Application US/09815242
```

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6698
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(951)
US-09-815-242-6698

alignment_scores:
  Quality: 80.50      Length: 218
  Ratio: 0.782       Gaps: 9
  Percent Similarity: 47.248      Percent Identity: 20.642

alignment_block:
US-09-215-435-225_copy_23_227 x US-09-815-242-6698 ..
Align seg 1/1 to: US-09-815-242-6698 from: 1 to: 951

11  GluAlaLeuLeuAspGluAspThrLeuPheCysGlnGlyLeuGluValPh 27
|||||
358 GAAGCCTTACTAGCAAGACACTGTTGTAATCATGCTGCGCCAGATATCCGTA 407
;
27  eTyrProGluLeuGlyAsnIleGlyCysLysValValProAspCysAsnA 44
|||||
408 TGAATATGATTAGTCAATTCGCTGCGCGCCGCGCCAGATATCCGTA 457
;
44  snTyrArgGlnLysIleThrSerTrpMetGluProIleValLysPhePro 60
|||||
458 GCTTCGCTGAG...TTACCACAATGGATTCCG..... 486

61  GlyAlaValAspGlyAlaThrTyrlleLeuValMetValAspProAspAl 77
486 .....

77  aProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValT 94
|||||
487 .....GAGAACAAAGACAATTTATGGATAAAAAAATTGTTA 523

94  hrAspIleLysGlyAlaAspLeuLysGlyLysIleGlnGlyGlnGlu 110
|||||
524 CCTATTGTACTGCGGG...ATTGCTGTGAAAAAATTTCTGCTGTTA 570
```

```

111  LeuSerAlaTyrGlnAlaProSerProAlaHisSerGlyPheHisAr 127
|||||
571 TTAAAGAAGAGATTTCAGATGTTGCTCAATTCATGCTGATCGCCAA 620

127  gTyr.....GlnPhePheValTyrLeuGlnGluGlyLysVal. 139
|||||
621 CTATGGAAAAATCCAGAAACACGTCGCGAACTTTGGACGCGAAATGT 670

140 .....IleSerLeuLeuProLysGluAsnLys 148
|||||
671 ATGCTCTTTGATGACCAATCAGTGTGAAATTAATCATGTTGATAAAAA 720

149  ThrArgGly...SerTrp.....LysMetAspArgPheLeuAs 160
|||||
721 GTATTGGGAAGACTGGTTTGATGGACACACTTGGCAAGCTACATTAA 770

160  nArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnA 177
|||||
771 C.....TGTGCAAAACCCAGATGTAATCGTCAAAATCTTAACCTCAG 811

177  snTyrGlnAspSerProThrLeu.....Gln 185
|||||
812 AGAAAAATGAACATAAACATTTAGTGGCTGCTCATTAGAATGTAGCCAG 861

186  AlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAlaGluIl 202
|||||
862 CATCTGCGCAACCGTTATGTAAAAAATAATTAATTAACAGACAGAAGT 911

202  eAla 203
|||||
912 TGCT 915
```

seq_name: /cgnl_7/ptodata/1/pna/US09_NEW_COMB.seq:US-09-922-340-6710

seq_documentation_block:

; Sequence 6710, Application US/09922340

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained

; FILE REFERENCE: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain

; CURRENT APPLICATION NUMBER: US/09/922,340

; CURRENT FILING DATE: 2001-08-03

; PRIOR APPLICATION NUMBER: 09/353,690

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: US 09/004,182

; NUMBER OF SEQ ID NOS: 12181

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6710

; LENGTH: 439

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-922-340-6710

alignment_scores:

Quality: 79.00 Length: 47

Ratio: 3.038 Gaps: 1

Percent Similarity: 55.319 Percent Identity: 36.170

alignment_block:

US-09-215-435-225_copy_23_227 x US-09-922-340-6710 ..

Align seg 1/1 to: US-09-922-340-6710 from: 1 to: 439

90 HisTrpLeuValThrAspIleLysGlyAlaAspLeuLysGlyLysIl 106

|||||

12 CACTGGCTGCTAACCAACATCCCGGTAAACCGGGTG.....GC 49

106 eGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProAlaHis 123

|||||


```

alignment_scores:      Length: 130
                        Quality: 78.50
                        Ratio: 1.246
                        Gaps: 5
Percent Similarity:    48.462   Percent Identity: 22.308

alignment_block:
US-09-215-435-225_COPY_23_227 x US-09-897-516-3030/rev ..

Align seg 1/1 to reverse of: US-09-897-516-3030 from: 1 to: 24961

        67 ThrTrrIleLeuValMetValAspProAspAlaProSerArgAlaGluPr 83
           ::::: ::::: |||||::::: :::::
22653 AGTTTGCGGTACAGGTTATGATCCTGCATGCCCAACGGGCAGT.... 22609

        83 oArgGlnArgPheTrpArgHisTrpLeuValThrAspIle..... 96
           ||||| |||||::::: |||||
22608 .....GGTTTTCTGG...CATTCGGGTGGCATTTGATATTCCGCTGAA 22569

        97 .....LysGlyAlaAspLeuLys 102
           |||||::::: |||

22568 TACAGCGTTTATCCCCGATGCCGACGGAGTGATGCGACAATATCACG 22519

103 LysGlyLysIleGlnGlyGlnGlu.....LeuSerAlaTy rGlnAl 116
    ||| |||||::::: ||||| ::::: :::::
22518 GCTGGCATATTACAGCAGAGAATAATGATTACGGCGCAGTTGTGTTGCGTGG 22469

116 aProSerProAlaHisSerGlyPheHisArgTy rGlnPhePheValt 133
    : ||||| :::: |||||::::: ||| |||
22468 AGCCTGCCCGCCAGAGGAGATAAAGCCCATCGTATATTTTCACGGTAC 22419

133 yrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThr 149
    : ::::: :::: |||||:::::
22418 ATGCCATGTGCGTAGAAAACACTGGGTATT..... 22390

150 ArgGlySerTriplysMetAspArgPheLeuAsnArgPheHisLeuGlyGl 166
    :: ::::: :::: |||||::::: |||
22389 .....AACGCTGAATACAAATCGGTCACACGTTTATGATCATCAGC 22346

166 uproGluAlaSerThrGlnPheMetThrGlnAsnTy rGln 179
    ::::: :::: |||||::::: |||||
22345 CATATGCTGGCAACCGCCCACTATTACAGGGCTATTACCAG 22306

```

THIS PAGE BLANK (USPTO)


```

251 TTGGCTGCAAGGTTGTCCTGATTGTAACAACATACAGACAGAAGATCAC 300
51  SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAla 67
301 TCCTGGATGAGCGATAGTCAAGATTCCCGGGGGCGGTGGACGGCCAAC 350
67  rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
351 CTATATCTCGTGTGATGGTATCCAGATGCCCTAGCAGACAGAACCCA 400
84  rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
401 GACAGAGATTCTGGAGACATTGGCTGTTAAGATATCAAGGGCGCGAC 450
101 LeuLysLysGlyLysIleGlnGlyGlnLeuLeuSerAlaTyrGlnAlaPr 117
451 CTGAAGGAAGGAAGATTCAAGGGCCAGGATTTACAGCTTACCAGGCTCC 500
117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPheValTyrL 134
501 CTCCCCAAGCGCACACAGTGGCTTCCATGCTACCACTTCTTGTCTATC 550
134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrAr 150
551 TTCAGGAAGGAACAGTCTCTCTCTCCCAAGGAAGAAACAAACTCG 600
150 gGlySerTrpLysMetAspArg...PheLeuAsnArgPheHisLeuGly 166
601 AGGCTCTTGGACAATGCCACAGATTCTTGAACCGGTTCACCTGGGCG 650
166 luProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSer. 181
651 AACCTGAAGCAAGCAAGCAAGTTCATGACCCAGCAACTTACCAAGACTC 700
182 ProThrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLys 197
701 CCAACCCCTCCAGGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 748

seq_name: gb_est93:BF854481

seq_documentation_block:
LOCUS BF854481 564 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-271200-005-d12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF854481
VERSION BF854481.1 GI:12242225
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&st2=MR2-EN0093-

```

```

271200-005-d12&st3=2000-12-27&st4=1)
Seq primer: puc 18 forward
High quality sequence stop: 562.
Location/Qualifiers
1..564
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 149 a 156 c 141 g 118 t
ORIGIN

alignment_scores:
Quality: 918.00 Length: 173
Ratio: 5.368 Gaps: 1
Percent Similarity: 98.844 Percent Identity: 98.844
alignment_block:
US-09-215-435-225_COPY_23_227 x BF854481
Align seg 1/1 to: BF854481 from: 1 to: 564
5 AsnSerProCysAlaHisGluAlaLeuLeuAspGluAspThrLeuPheCys 21
46 AACAGCCCGTGTGCCCATCAGGCCCTCTTGGACGAGGACACCCCTCTTTG 95
21 sGlnGlyLeuGluValPheTyrProGluLeuGlyAsnIleGlyCysLysV 38
96 CCAGGGCCTTGAAGTTTCTACCCAGAGTGGGGAACATTCCTGCAAGG 145
38 alValProAspCysAsnAsnTyrArgGlnLysIleThrSerTrpMetGlu 54
146 TTGTTCTCTGATTGTAACAACACTACAGACAGACATCACCTCTGGATGGAG 195
55 ProIleValLysPheProGlyAlaValAspGlyAlaThrTyrIleLeuVa 71
196 CCATATAGTCAAGTTCCGGGGCGCGTGGACGGCAACCTATATCTCTGT 245
71 lMetValAspProAspAlaProSerArgAlaGluProArgGlnArgPhe 87
246 GATGGTGGATCCAGATGCCCTTAGCAGACAGACCCAGACAGATTC 295
88 TrpArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysGly 104
296 TGGAGACATTGGCTGGTAAACAGATATCAAGGGCGCGACCTGAAGAAGG 345
104 yLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerProProA 121
346 GAAGATTTCAGGGCGCAGAGATTATCAGCCTACCAAGGCTCCCTCCACC 395
121 laHisSerGlyPheHisArgTyrGlnPheValTyrLeuGlnGluGly 137
396 CACACAGTGGCTCCATCGCTACCAAGTTCTTCTATCTTATCTTACGAAGA 445
138 LysValIleSerLeuLeuProLysGluAsnLysThrArgGlySerTrp 154
446 AAAGTCATCTCTCTCTCCCAAGGAACAAACATCGAGGCTCTTGGAA 495
154 sMetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAlaSer 171
496 AATGCACAGATTCTGAAACCTTCCACCTGGGCGAAGAACTGAAGCAAGCA 545
171 hrGlnPheMetThrGln 176
546 CCCAGTTTCATGACCCAG 562

```

seq_name: gb_est93:BF852989

seq_documentation_block:
 LOCUS BF852989 603 bp mRNA EST 16-JAN-2001
 DEFINITION MR2-EN0093-191200-001-al2 EN0093 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF852989
 VERSION BF852989.1 GI:12240837
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 603)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
 COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
 191200-001-al2&t3=2000-12-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 560.

FEATURES

source

1..603
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0093"
 /dev_stage="Adult"
 /note="Organ: lung_normal; Vector: puc18; Site: 1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 140 a 157 c 171 g 135 t

alignment_scores:

Quality: 775.00 Length: 149
 Ratio: 5.345 Gaps: 0
 Percent Similarity: 97.315 Percent Identity: 96.644

alignment_block:

US-09-215-435-225_COPY_23_227 x BF852989 ..

Align seg 1/1 to: BF852989 from: 1 to: 603

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17

157 GACGAGGATGAGAACAGCCGCTGTGCCCATGAGGCCCTCTTGAACGAGGA 206

17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34

207 CACCCTCTTTCCAGGGCCCTGAAGTCTTCTACCCAGAGTTGGGGAACA 256

34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
 257 TTGGCTGCAAGGTTGCTCTGATTGTAACACTACACAGACAGATCACC 306
 51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
 307 TCCTGTGATGAGCGCATAGTCAAGTTCCCGGGGGCGGTGGACGGCGCAAC 356
 67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
 357 CTATATCCCTGGGTATGGTGGATCCAGATCCCTACGACAGCAACCCCA 406
 84 rGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
 407 GACAGAGATTCCTGGACATTCGGCTGGTAACAGATATCAAGGGCGCCGAC 456
 101 LeuLysLysGlyLysIleGlnGlnGluLeuSerAlaTyrGlnAlaPr 117
 457 CTGAAGAAAGGGAAGATTTCAGGGCGCAGGAGTTATCAGCTACACGGCTCC 506
 117 oSerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
 507 CTTCCCGCCCGGCACACAGTGGCTTCCATCCTAACAGTTCTTTGTCTATC 556
 134 euGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThr 149
 557 TTCAGGAGGGAAGTCATCTCTCTCTTCCCAAGGGAACAACT 603

seq_name: gb_est93:BF852995

seq_documentation_block:

LOCUS BF852995 563 bp mRNA EST 16-JAN-2001
 DEFINITION MR2-EN0093-191200-001-e01 EN0093 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF852995
 VERSION BF852995.1 GI:12240843
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
 191200-001-e01&t3=2000-12-19&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 563.

Location/Qualifiers

1..563

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="EN0093"

/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site: 1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 126 a 148 c 164 g 125 t
ORIGIN

alignment_scores:
Quality: 744.00 Length: 135
Ratio: 5.511 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-215-435-225_COPY_23_227 x BF852995 ..

Align seg 1/1 to: BF852995 from: 1 to: 563

```

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
157 GACGAGGATGAGAACAGCCGCTGTGCCCATGAGGCCCTCTTGGACGAGGA 206
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
207 CACCCCTCTTTGCCAGGCCCTTGGAAGTTTCTACCCAGAGTTGGGGACAA 256
34 leglyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
257 TTGGCTGCAAGGTTTCTCCTGATTGTAACAACATACAGACAGAGATCACC 306
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
307 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCCCTGGACGCGCCAAC 356
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
357 CTATATCTCTGTGTGGTGGATCCAGATGCCCTAGCAGAGCAGAACCCA 406
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
407 GACAGAGATCTCGAGAGATTGGCTGTGTACAGATATCAAGGGCGCCGAC 456
101 LeuLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaPr 117
457 CTGAAGAAGGAAGATTACAGGGCAGGAGTTATCAGCTACAGGCTCC 506
117 oSerProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrL 134
507 CTCGCCACGGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTTGTCTATC 556
134 euGln 135
557 TTCAG 561

```

seq_name: gb_est93:BF853000

seq_documentation_block:
LOCUS BF853000 558 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-f12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853000
VERSION BF853000.1 GI:12240848
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 558)
Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R., Reis, L.F., de Souza, S.J. and

TITLE
JOURNAL
MEDLINE
COMMENT

Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-f12&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 558.

FEATURES
source

1..558
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 124 a 146 c 163 g 124 t 1 others
ORIGIN

alignment_scores:
Quality: 727.00 Length: 133
Ratio: 5.508 Gaps: 0
Percent Similarity: 99.248 Percent Identity: 99.248

alignment_block:
US-09-215-435-225_COPY_23_227 x BF853000 ..

Align seg 1/1 to: BF853000 from: 1 to: 558

```

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
159 GACGAGGATGAGAACAGCCGCTGTGCCCATGAGGCCCTCTTGGACGAGGA 208
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
209 CACCCCTCTTTGCCAGGCCCTTGAAGTTTCTTACCCAGAGTTGGGGAAACA 258
34 leglyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
259 TTGGCTGCAAGGTTTCTCCTGATTGTAACAACATACAGACAGAGATCACC 308
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
309 TCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGCCCTGGACGCGCCAAC 358
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
359 CTATATCTCTGTGTGGTGGATCCAGATGCCCTTAGCAGAGCAGAACCCA 408
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
409 GACAGAGATCTCGAGAGATTGGCTGTGTACAGATATCAAGGGCGCCGAC 458
101 LeuLysLysGlyLysIleGlnGlnLeuSerAlaTyrGlnAlaPr 117
459 CTGAAGAAGGAAGATTACAGGGCAGGAGTTATCAGCTACAGGCTCC 508

```

117 oSerProAlaHisSerGlyPheHisArgTyrGlnPheValTyr 133
||||| |||||
509 CTCCCAGCGGCACACAGTGCGTTCATCGCTACCAGTTCTTTGTCTAT 557

seq_name: qb_est25:AI816715

```

seq_documentation_block:
LOCUS      A1816715      547 bp      mRNA      21-DEC-1999
DEFINITION wJ43b08.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:240559 3,
            similar to SW:D2_ONCVO P54187 D2 PROTEIN ; , mRNA sequence.
ACCESSION  A1816715
VERSION    A1816715.1 GI:5435794
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 547)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: M. Bento Soares, Ph.D.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Insert length: 903 Std Error: 0.00
            Seq primer: -400P from Gibco
            High quality sequence stop: 453.
            Location/Qualifiers
                1..547
FEATURES
source

```

FACSIMILE DONATION.			
BASE COUNT	102 a	131 c	160 g
ORIGIN			
			154 t

alignment_scores:		
Quality:	718.00	Length: 141
Ratio:	5.129	Gaps: 0
Percent Similarity:	99.291	Percent Identity: 98.582

alignment_block:
US-09-215-435-225_COPY_23_227 x AI816715/rev

Align seg 1/1 to reverse of: AI816715 from: 1 to: 547

65 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 81
|||||
547 GGCGCAACTATATCTGGTGATGGATCCAGATGCCCTAGCAGACG 491

81	agLupProArgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysG	98
497	AGAACCCAGACAGAGATCTCGAGACATTTGGCTGGTAACAGATATCAAG	448
98	lyAlaAspLeuLysGlySlyLysIleGlnGlyGlnLeuSerAlaTyr	114
447	CGCGGAGCCTGAAGGAGGGAAGATTGAGGCCAGGAGTATCAGCCTAC	398
115	GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePh	131
397	CAGGCTCCCTCCCAACCCGACACAGATGGCTTCCATCGCTACCAAGTCTT	348
131	eValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnL	148
347	TGCTATCTTCAGGAAGGAAAGATCATCTCTCTCTCCCAAGGAAACA	298
148	ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu	164
297	AAACTCGAGGCTCTTGGAAAATGACAGATTTTGAACCGCTTCCACTTG	248
165	GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspse	181
247	GCGAAACCTGAAGCAAGCACCCAGTTCATGACCCAGAACTACCAAGACTC	198
181	rProThrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysA	198
197	ACCAACCTCCAGGCTCCAGAGAAAGGCGGAGCCAGCCCAAGCACAA.A	149
198	snGlnAlaGluIleAlaAlaCys	205
148	ACCAGGCGGAGATAGTGCCTGC	126
seq_name: gb_est25:AI857902		
seq_documentation_block:		
LOCUS	AI857902	547 bp mRNA EST 21-DEC-1999
DEFINITION	wJ68a01.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407944 3'	
	similar to SW:02_ONCVO P54187 D2 PROTEIN ;, mRNA sequence.	
ACCESSION	AI857902	
VERSION	AI857902.1	GI:5511518
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.	
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
	Unpublished (1997)	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgaps@remail.nih.gov	
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.	
	CDNA Library preparation: M. Bento Soares, Ph.D.	
	CDNA Library Arrayed by: Greg Lennon, Ph.D.	
	DNA Sequencing by: Washington University Genome Sequencing Center	
	Clone distribution: NCI-CGAP clone distribution information can	
	found through the I.M.A.G.E. Consortium/LLNL at:	
	www-bio.llnl.gov/bbrp/image/image.html	
	Insert Length: 752 Std Error: 0.00	
	Seq primer: -400P from Gibco	
	High quality sequence stop: 450.	
FEATURES	Location/Qualifiers	
source	1..547	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:2407944"	
	/clone_lib="NCI-CGAP_Lu19"	
	/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"	

FEATURES SOURCE

```

/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- Oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 102 a 130 c 160 g 155 t
ORIGIN

```

```

alignment_scores:
  Quality: 718.00      Length: 141
  Ratio: 5.129        Gaps: 0
  Percent Similarity: 99.291      Percent Identity: 98.582

alignment_block:
US-09-215-435-225_COPY_23_227 x AI857902/rev
Align seg 1/1 to reverse of: AI857902 from: 1 to: 547

```

```

65 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 81
547 GCGCAACCTATATCTTGGTGGATCCAGATGCCCTAGCAGAGC 498
81 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspLeuLysG 98
497 AGAACCCAGACAGAGATCTGGAGACATTTGGTGGTAAACAGATATCAGG 448
98 lYAlaAspLeuLysGlyLysIleGlnGlnGluLeuSerAlaTyr 114
447 GCGCGACCTGAAGAGGAAAGATTCAGGCGCAGGAGTTATCAGCCTAC 398
115 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePh 131
397 CAGGCTCCCTCCCGCCAGGACAGTGGCTTCCATCGCTACCAAGTTCTT 348
131 eValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnL 148
347 TGTCTATCTTCAGGAGGAAAGATCATCTCTCTCTCCCAAGGAAACAA 298
148 ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu 164
297 AAACCTCGAGGCTCTTGGAAAAATGGACAGATTTCTCAACCGCTTCCAC 248
165 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspse 181
247 GCGCAACCTGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGACTC 198
181 rProThrLeuGlnAlaProArgGluArgAlaSerGluProLysHisLysA 198
197 ACCAACCTCCAGGCTCCAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAG 149
198 snglAlaGluIleAlaLacys 205
148 ACCAGGCGGAGATAGCTGCCTGC 126
seq_name: gb_est69:BE044451

```

```

seq_documentation_block:
LOCUS BE044451 523 bp mRNA EST 08-JUN-2000
DEFINITION ho45g08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:3040382 3' similar to SW:DI_ONCVO P54186 DI PROTEIN ;, mRNA
sequence.
ACCESSION BE044451
VERSION BE044451.1 GI:8361504
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
1 (bases 1 to 523)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapps@mail.nih.gov
This clone is available royalty-free through LLM: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 465.

```

```

FEATURES
  source
    1..523
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:3040382"
      /lab_host="DH10B"
      /lab_host="DH10B"
      /notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 96 a 128 c 159 g 140 t
ORIGIN

```

```

alignment_scores:
  Quality: 709.00      Length: 141
  Ratio: 5.101        Gaps: 0
  Percent Similarity: 98.582      Percent Identity: 97.163

alignment_block:
US-09-215-435-225_COPY_23_227 x BE044451/rev
Align seg 1/1 to reverse of: BE044451 from: 1 to: 523

```

```

65 GlyAlaThrTyrIleLeuValMetValAspProAspAlaProSerArgAl 81
523 GCGCAACCTATATCTTGGTGGATCCAGATGCCCTAGCAGAGC 474
81 aGluProArgGlnArgPheTrpArgHisTrpLeuValThrAspLeuLysG 98
473 AGAACCCAGACAGAGATCTGGAGACATTTGGTGGTAAACAGATATCAAG 424
98 lYAlaAspLeuLysGlyLysIleGlnGlnGluLeuSerAlaTyr 114
423 GCGCGACCTGAAGAGGAAAGATTCAGGCGCAGGAGTTATCAGCCTAC 374
115 GlnAlaProSerProProAlaHisSerGlyPheHisArgTyrGlnPhePh 131
373 CAGGCTCCCTCCCGCCAGGACAGTGGCTTCCATCGCTACCAAGTTCTT 324
131 eValTyrLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnL 148
323 TGTCTATCTTCAGGAGGAAAGATCATCTCTCTCTCCCAAGGAAACAA 274
148 ysThrArgGlySerTrpLysMetAspArgPheLeuAsnArgPheHisLeu 164
273 AAACCTCGAGGCTCTTGGAAAAATGGACAGATTTCTGAACCGCTTCCAC 224
165 GlyGluProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspse 181
223 GCGCAACCTGAAGCAAGCACCAGTTCATGACCCAGAACTACCAAGACTC 174

```


181 rProThrLeuGlnAlaProArgGluAlaSerGluProLysHisLysA 198
 |||||
 173 ACCAACCTCCAGGCTCCAGAGGAGGCCAGGAGCCCAAGCACAA.A 125
 |||||
 198 snGlnAlaGluLeuAlaAlaCys 205
 |||||
 124 ACCAGCGGAGATAGCTGCCTGC 102

seq_name: gb_est17:AI218954

seq_documentation_block: 552 bp mRNA EST 21-DEC-1998
 LOCUS AI218954
 DEFINITION qg72908.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:1840766 3' similar to SW:BBP_RAT_P31044
 PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN ; mRNA sequence.
 ACCESSION AI218954
 VERSION AI218954.1 GI:3801157
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 552)
 AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 This clone is available royalty-free through LILN ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1054 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 423.

Location/Qualifiers
 1..552

FEATURES
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1840766"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI_CGAP_GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 98 a 129 c 158 g 167 t
 ORIGIN

alignment_scores:
 Quality: 699.00 Length: 138
 Ratio: 5.216 Gaps: 0
 Percent Similarity: 97.101 Percent Identity: 96.377

alignment_block:

US-09-215-435-225_COPY_23_227 x AI218954/rev ..

Align seg 1/1 to reverse of: AI218954 from: 1 to: 552

68 TvrIleuValMetValAspProAspAlaProSerArgAlaGluProAr 84

552 TATATCTGGTGATGATCCAGTACCGCCCTAGCAGACCAAGCCAG 503

84 gGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspL 101

|||||
 502 GCAGAGATTGGAGACCATTGGCTGTGTACAGATATCAAGGGCGCCGACC 453
 |||||
 101 euLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPro 117
 |||||
 452 TGAAGGAAGGAAGATTTCAGGCCAGAGTTATCAGCCTACCAAGCTCCC 403
 |||||
 118 SerProProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLe 134
 |||||
 402 TCCCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTCTCTATCT 353
 |||||
 134 uGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgG 151
 |||||
 352 TCAGGAAGGAAAAAGTCTCTCTCTCCATCCCAAGGAAAAACAACTCGAG 303
 |||||
 151 LySerTrpLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluPro 167
 |||||
 302 GCTCTTGGAAATGGACAGATTTCTGAACCGTTTCCACCTGGGGCAACCT 253
 |||||
 168 GluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrLe 184
 |||||
 252 GAAGCAAGCACCCAGTTCATGACCCAGAACTACCAAGACTCACCACCCCT 203
 |||||
 184 uGlnAlaProArgGluArgAlaSerGluProLysHisLysAsnGlnAlaG 201
 |||||
 202 CCAGGCTCCAGAGAAGGGCCAGCGAGCCCAAGCAACAAACAGCGCG 153
 |||||
 201 luIleAlaAlaCys 205
 |||||
 152 AGATAGCTGCCTGC 139

seq_name: gb_est26:AI936064

seq_documentation_block: 534 bp mRNA EST 08-MAR-2000
 LOCUS AI936064
 DEFINITION wo47hi2.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2458535 3'
 similar to SW:D2_ONCVO_P54187 D2 PROTEIN ; mRNA sequence.
 ACCESSION AI936064
 VERSION AI936064.1 GI:5674934
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 534)
 AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 942 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 398.

Location/Qualifiers

1..534

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2458535"

/clone_lib="NCI_CGAP_Gas4"

/tissue_type="poorly differentiated adenocarcinoma with

signed ring cell features"

/lab_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.69 kb. Life Technologies catalog #:

BASE COUNT 96 a 133 c 165 g 140 t

alignment_scores:
Quality: 695.00 Length: 145
Ratio: 4.964 Gaps: 0
Percent Similarity: 96.552 Percent Identity: 96.552

alignment_block:

US-09-215-435-225_COPY_23_227 x A1936064/rev ..

Align seg 1/1 to reverse of: A1936064 from: 1 to: 534

```

61 GlyAlaValAspGlyAlaThrTyrIleLeuValMetValAspProAspAl 77
|||||
531 GGGCCCGTGGCGGCAACCTATATCTCGTGGTGGTGGTGGTGGTGGTGGT 482
|||||
77 aProSerArgAlaGluProArgGlnArgPheTrpArgHisTrpLeuValT 94
|||||
481 CCTAGCAGACAGACACCCAGACAGAGATTCTGGAGACATTGGCTGGTAC 432
|||||
94 hrAspIleLysGlyAlaAspLeuLysGlyLysIleGlnGlu 110
|||||
431 AAGATAC.AAGGGCCCGGACCTGAAGAAAGAAAGATTTCAGGCGCAGGAG 383
|||||
111 LeuSerAlaTyrGlnAlaProSerProProAlaHisSerGlyPheHisAr 127
|||||
382 TTATCAGCCTACAGGCTCCCTCCCGCCAGCACAGTGGCTTCCATCG 333
|||||
127 gTyrGlnPhePheValTyrLeuGlnGluGlyLysValIleSerLeuLeup 144
|||||
332 CTACCAAGTCTTCTGCTATCTTCAGGAAGGAAAGTCTCTCTCTCTTC 283
|||||
144 rOLysGluAsnLysThrArgGlySerTrpLysMetAspArgPheLeuAsn 160
|||||
282 CCAGGAAACAACTCCAGGCTCTTGGAAATGGACAGATTCTCGAAC 233
|||||
161 ArgPheHisLeuGlyGluProGluAlaSerThrGlnPheMetThrGlnAs 177
|||||
232 CGCTTCCACCTGGCGGAACCTGAAGCAAGCACCCAGTTCATGACCCAGAA 183
|||||
177 nTyrGlnAspSerProThrLeuGlnAlaProArgGluArgAlaSerGluP 194
|||||
182 CTACCAAGGACTCACCAACCTCCAGGCTCCAGAGAGAGAGAGAGAGAGAG 133
|||||
194 rOLysHisLysAsnGlnAlaGluIleAlaAlaCys 205
|||||
132 CCAAGCACAA.AACCAGGCAGAGATAGTGCCTGC 99
|||||

```

seq_name: gb_est93:BF846558

seq_documentation_block: 532 bp mRNA EST 16-JAN-2001
LOCUS BF846558
DEFINITION OVO-EN0058-261000-441-a06 EN0058 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF846558
VERSION BF846558.1 GI:12233708

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 532)

REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Simpson, D. H., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE

JOURNAL
MEDLINE
COMMENT

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-EN0058-261000-441-a06&t3=2000-10-26&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 532.

Location/Qualifiers

FEATURES

1..532

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="EN0058"

/dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

140 a 158 c 138 g 96 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 687.50 Length: 136

Ratio: 5.169 Gaps: 1

Percent Similarity: 97.794 Percent Identity: 95.588

alignment_block:

US-09-215-435-225_COPY_23_227 x BF846558 ..

Align seg 1/1 to: BF846558 from: 1 to: 532

```

70 LeuValMetValAspProAspAlaProSerArgAlaGluProArgGlnAr 86
|||||
27 CTGGTGATGGTGGAT...CAGATGCCCTAGCAGACAGAACCCAGACAGAG 73
|||||
86 gPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAspLeuLysL 103
|||||
74 ATCTGGAGACATTGGCTGGTAAACAGATATCAAGGGCGCCGACCTGAAGG 123
|||||
103 ySGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaProSerPro 119
|||||
124 AAGGAGAGATTACGGCCAGGAGCTTATCAGCTTACAGGCTCCCTCCCA 173
|||||
120 ProAlaHisSerGlyPheHisArgTyrGlnPhePheValTyrLeuGlnG 136
|||||
174 CCGGCACACAGTGGCTTCCATCGCTACCAAGTCTCTTGTCTATCTTCAGGA 223
|||||
136 uGlyLysValIleSerLeuLeuProLysGluAsnLysThrArgGlySerT 153
|||||
224 AGNAAAGTCACCTCTCTCTCCCAAGGAAACAACTCGAGGCTCTT 273
|||||
153 rPLysMetAspArgPheLeuAsnArgPheHisLeuGlyGluProGluAla 169
|||||
274 GGAATATGGACAGACTTCTCAACCGTTTCCACCTGGCGCAACCTGAAGCA 323
|||||
170 SerThrGlnPheMetThrGlnAsnTyrGlnAspSerProThrLeuGlnAl 186
|||||
324 AGCACCAGTTCTATGACCCAGAACTACCAAGACTCACCACCTCCAGGC 373
|||||
186 aProArgGluArgAlaSerGluProLysHisLysAsnGlnAlaGluIleA 203
|||||

```

```

|||||
374 TCCAGAGAAAGGCGAGCGAGCCCAAGCACAAAACAGCGCGAGATAG.423
203 laAlaCys 205
|||||
424 CTGCGCTGC 431
seq_name: gb_est110:W32197

seq_documentation_block:
LOCUS W32197 451 bp mRNA EST 10-OCT-1996
DEFINITION W32197 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGE:321345 5' similar to SW:PBPL_RAT P31044 23 KD
MORPHINE-BINDING PROTEIN ;, mRNA sequence.
ACCESSION W32197
VERSION W32197.1 GI:1313189
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 705 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 429.
FEATURES
Location/Qualifiers
1..451
/organism="Homo sapiens"
/db_xref="GDB:1259003"
/db_xref="taxon:9606"
/clone="INAGE:321345"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site.1: Not I; Site.2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer
[5'-TGTTACCAATCTCAAGTGGAGCGCGCGACCAATTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."
BASE COUNT 124 a 122 c 109 g 90 t 6 others
ORIGIN

alignment_scores:
Quality: 686.00 Length: 152
Ratio: 4.667 Gaps: 2
Percent Similarity: 96.711 Percent Identity: 94.079

alignment_block:

```

```

US-09-215-435-225_COPY_23_227 x W32197 ..
Align seg 1/1 to: W32197 from: 1 to: 451

34 IleGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleTh 50
|||||
2 APTGGCTGCAAGGTGTCTCTGATTGTAACAACACTACAGACAGAAGATCAC 51
|||||
50 rSerTirpMetGluProIleValLysPheProGlyAlaValAspGlyAla 67
|||||
52 CTCTCGGATGAGCGGAGTAGTCAAGTTCCTCGGGGGCCGTGGACNG.CANA 100
|||||
67 hrTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluPro 83
|||||
101 CCTATATCTCGTGATGCTGGTGGATCCAGATGCCCTAGCAGANA.GAACCC 149
|||||
84 ArgGlnArgPheTirpArgHisTirpLeuValThrAspIleLysGlyAla 100
|||||
150 AGACAGAGATCTGGAGACATTGGCTGGTAACAGATATCAAGGGCGCCGA 199
|||||
100 pLeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlap 117
|||||
200 CTGAGGAGGAGGAGATTCAGGGCCAGGAGTTATCAGCCTACCAGNT.C 248
|||||
117 rSerProProAlaHisSerGlyPheHisArgTyrGln.PhePheValTy 133
|||||
249 CCTCCCCACCGGCACACAGTGGCTTCATCGCTACCAAGTCTCTTGTCTA 298
|||||
133 rLeuGlnGluGlyLysValIleSerLeuLeuProLysGluAsnLysThr. 149
|||||
299 TCTTCAGGAGGAGGAGGATCATCTCTCTCTCCCAAGGAAAAACAACATC 348
|||||
150 ArgGlySerTirpLysMetAspArgPheLeuAsnArgPheHisLeuGly 166
|||||
349 NGAGGCTCTTGGAAATGACAGATTTCTGAACCGTTCCACCTGGGCGA 398
|||||
166 uProGluAlaSerThrGlnPheMetThrGlnAsnTyrGlnAspSerProt 183
|||||
399 ACCTGAAGCAAGCACCCAGTTTCATGACCCAGAACTTACCAGGACTCANCA 448
|||||
183 hr 183
||
449 CC 450
seq_name: gb_est93:BF853920

seq_documentation_block:
LOCUS BF853920 499 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-261200-004-d01 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853920
VERSION BF853920.1 GI:12241664
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 499)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

```

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl1-MR2&tl2-MR2-EN0093-261200-004-d01&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 227.
Location/Qualifiers
1. 499
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 119 a 129 c 148 g 103 t
ORIGIN

alignment_scores:
Quality: 683.00 Length: 127
Ratio: 5.421 Gaps: 0
Percent Similarity: 99.213 Percent Identity: 97.638
alignment_block:
US-09-215-435-225_COPY_23_227 x BF853920 ..
Align seg 1/1 to: BF853920 from: 1 to: 499
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
114 GACGAGGATGAGAACACCGCGTGTGCCATGAGGCCCTCTTGGACGAGGA 163
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
164 CACCCCTCTTTTGGCAGGCGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 213
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
214 TTGCTCACAAGTTTCTTCTGATGTGTACAACTACACAGACAGATCACC 263
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67
|||||
264 TCCTGGATGGAGCGCATAGTCAAGTCCCGGGCGCGTGGACGGCGCAA 313
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
:|||||
314 CTATATCTCTGTGTGTGTGTATCCAGATGCCCTAGCAGACAGACCCA 363
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
|||||
364 GACAGACATCTCGAGACATGTGGTGTACAGATATCAAGGGCGCGAC 413
101 LeuLysLysGlyLysIleGlnGlnGlnLeuSerAlaTyrGlnAlaPr 117
|||||
414 CTGAAGAAAGGAAGATTTCAGGGCCAGGAGTTATCAGGCTACCAAGGCTCC 463
117 oSerProAlaHisSerGlyPheHisArg 127
|||||
464 CTCGCCACCGGCACACAGTGGCTTCCATCGC 494

seq_name: gb_est93:BF846063
seq_documentation_block:
LOCUS BF846063 464 bp mRNA EST 16-JAN-2001
DEFINITION QV0-EN0055-181000-439-flt EN0055 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF846063
VERSION BF846063.1 GI:12233213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 464)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl1-QV0&tl2-QV0-EN0055-
181000-439-flt&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 464.
Location/Qualifiers
1..464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0055"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 134 c 122 g 114 t
ORIGIN

alignment_scores:
Quality: 681.00 Length: 124
Ratio: 5.492 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-215-435-225_COPY_23_227 x BF846063/rev ..
Align seg 1/1 to reverse of: BF846063 from: 1 to: 464
1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
|||||
390 GACGAGGATGAGAACACCGCGTGTGCCATGAGGCCCTCTTGGACGAGGA 341
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
|||||
340 CACCCCTCTTTTGGCAGGCGCTTGAAGTTCTTACCCAGAGTTGGGGAACA 291
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
|||||
290 TTGGCTGCAAGTTGTCTCTGATTTGAACAACACTACACAGACAGATCACC 241
51 SerTrpMetGluProIleValLysPheProGlyAlaValAspGlyAlaTh 67

alignment_scores:
 Quality: 676.00 Length: 126
 Ratio: 5.408 Gaps: 0
 Percent Similarity: 99.206 Percent Identity: 97.619

alignment_block:
 US-09-215-435-225_COPY_23_227 x BF853928 ..
 Align seg 1/1 to: BF853928 from: 1 to: 536

```

1 AspGluAspGluAsnSerProCysAlaHisGluAlaLeuLeuAspGluAs 17
157 GAGGAGGATGAGAACAGCCCGTGTGCCCATGAGGCCCTCTTGGACGAGGA 206
17 pThrLeuPheCysGlnGlyLeuGluValPheTyrProGluLeuGlyAsnI 34
207 CACCCTCTTTTGGCCAGGGCTTGAAGTTTCTACCCAGAGTTGGGGAACA 256
34 leGlyCysLysValValProAspCysAsnAsnTyrArgGlnLysIleThr 50
257 TTGGCTGCAAGGTCGTTCTCTGATTGTAACAACCTACAGACAGAAGATCACC 306
51 SerTrpMetGluProIleValLysPheProGlyAlaValAlaAspGlyAlaTh 67
307 TCTTGGATGGAGCCGATAGTCAAAGTTCCCGGGGGCGGTGGACGGCGCAAC 356
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
357 CTATATCTCTGGTGTGGTGGATCCAGATCCCTTACAGAGCAGACAGACCCA 406
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
407 GACAGAGATTCTCGAGACATTGGCTGTACACAGATATCAAGGGCGCGCAC 456
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
457 CTGAAGAAAGGGGAAGATTTCAGGGGCGAGGAGTTATCAGGCTTAACAGGGTCC 506
117 oSerProAlaHisSerGlyPheHis 126
507 CTCCTCCACCGGCACACAGTGGCTTCCAT 534

```

```

240 TCTGTGATGAGCCGATAGTCAAGTTCCGGGGCGCGTGGACGGCGCAAC 191
67 rTyrIleLeuValMetValAspProAspAlaProSerArgAlaGluProA 84
190 CTATATCTCTGGTGTGGTGGATCCAGATCCCTTACAGAGCAGAACCCA 141
84 rgGlnArgPheTrpArgHisTrpLeuValThrAspIleLysGlyAlaAsp 100
140 GACAGAGATTCTGGACACATTGGCTGTAAACAGATATCAAGGGCGCGCAC 91
101 LeuLysLysGlyLysIleGlnGlyGlnGluLeuSerAlaTyrGlnAlaPr 117
90 CTGAAGAAAGGGGAAGATTTCAGGGGCGAGGAGTTATCAGGCTTACAGGCTCC 41
117 oSerProAlaHisSerGly 124
40 CTCCTCCACCGGCACACAGTGGC 19
seq_name: gb_est93:BF853928

seq_documentation_block:
LOCUS BF853928 536 bp mRNA EST 16-JAN-2001
DEFINITION MR2-EN0093-261200-004-g12 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853928
VERSION BF853928.1 GI:12241672
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 536)
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR2&t2=MR2-EN0093-
261200-004-g12&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 497.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0093"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
121 a 139 c 164 g 112 t
BASE COUNT
ORIGIN

```

